

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 00:52:36 ; Search time 108 Seconds
(without alignments)
8101.378 Million cell updates/sec

Title: US-09-965-631-3
Perfect score: 2853
Sequence: 1 atgcttctgtctggcctact.....gcgctctcaggccgtgctga 2853

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 24

Total number of hits satisfying chosen parameters: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database :
Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description

No matches found

Search completed: May 16, 2003, 04:04:23
Job time : 108 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 00:46:46 ; Search time 3592 Seconds
(without alignments)
12863.509 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtctctgtggcctcct.....ggtctctgagcgctgtctga 2853

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 24

Total number of hits satisfying chosen parameters: 15

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_fod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	53	1.9	544	12 BF078689	BF078689 229090 MA
C 2	39	1.4	195	10 AW935807	AW935807 QV3-DT001
C 3	38	1.3	528	17 AQ809642	AQ809642 HS-4758.A
C 4	36	1.3	961	10 BB612189	BB612189 BB612189
C 5	35	1.2	469	17 AZ842795	AZ842795 ZM0141119
C 6	35	1.2	522	10 BE553572	BE553572 ur47b04.y

7	30	1.1	305	12	BF320986	BF320986 uz57h10.y
8	29	1.0	400	17	AQ440250	AQ440250 HS-5071.B
9	28	1.0	415	10	AW485848	AW485848 68315 MAR
10	27	0.9	169	17	B01986	B01986 CSRL-145A3-
11	26	0.9	627	10	BE281680	BE281680 601099657
12	25	0.9	247	10	BB604671	BB604671 BB604671
13	25	0.9	849	13	BI103177	BI103177 602889417
14	24	0.8	464	10	BE666088	BE666088 149119 MA
15	24	0.8	664	13	BI692986	BI692986 603344521

ALIGNMENTS

RESULT 1
LOCUS BF078689/c
DEFINITION 229090 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF078689
VERSION BF078689.1 GI:10872519
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 544)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Langreid, W.W.
and Keele, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 43 row: P column: 17
Seq primer: ATTAGGTGACACTATAG.
FEATURES
Location/Qualifiers
1..544
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 117 a 154 c 169 g 104 t
ORIGIN
Query Match 1.9%; Score 53; DB 12; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 ACGGGCGGCGACTCTACCGCATCCAGCATCTCTCAACCCCATCATCGT 794
|||||
DB 361 ACGGGCGGCGACTCTACCGCATCCAGCATCTCTCAACCCCATCATCGT 309

RESULT 2
LOCUS AW935807/c
DEFINITION QV3-DT0019-081299-039-c08 DT0019 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW935807

```

VERSION      AW935807.1  GI:81111213
KEYWORDS     EST.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE    1 (bases 1 to 195)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
             Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
             Goldman,G.H., Cavallho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
             Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
             M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
TITLE        Shotgun sequencing of the human transcriptome with ORF expressed
             sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-27047001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the PAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-QV3-DT0019-081
             299-039-c08&ts=1999-12-08&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 58
             High quality sequence stop: 109.
FEATURES     source
             1..195
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="DR0019"
             /dev_stage="Adult"
             /note="Organ: denis_drash; Vector: puc18; Site_1: SmaI;
             Site_2: SmaI; A mini-library was made by cloning products
             derived from ORESTES PCR (U.S. Letters Patent application
             No. 196,716 - Ludwig Institute for Cancer Research)
             profiles into the puc 18 vector. Reverse transcription of
             tissue mRNA and cDNA amplification were performed under
             low stringency conditions."
BASE COUNT   37 a 60 c 52 g 46 t
ORIGIN
Query Match 1.4%; Score 39; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1542 GTGGGATGTTCTCTGGCCCAATGGATCCCTATGGCCC 1580
      |||||||||||||||||||||||||||||||||||||||
Db 155 GGTGGATGTTCTCTGGCCCAATGGATCCCTATGGCCC 117

RESULT 3
A0809642
LOCUS       HS_4758_Al_D11_77A CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=4758 Col=21 Row=G, DNA sequence.
ACCESSION  A0809642
VERSION     A0809642.1 GI:5728884
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 528)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.

```

```

TITLE        Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE      99380589
COMMENT      Contact: Mahairas G.G., Wallace J.C., Hood L
             High throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallaceu.washington.edu
             Clones may be purchased from Research Genetics (info@resgen.com).
             BAC end Web Server: http://www.htsc.washington.edu
             Plate: 4758 row: G column: 21
             Seq primer: T7
             Class: BAC ends
             High quality sequence stop: 528.
FEATURES     source
             1..528
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="CIT Approved Human Genomic Sperm Library D"
             /sex="male"
             /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
             E-Coli DH10B"
BASE COUNT   138 a 145 c 100 g 142 t 3 others
ORIGIN
Query Match 1.3%; Score 38; DB 17; Length 528;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2114 GGCCTCAAGCATCGACATCGCCGCGGCTTACAAA 2151
      |||||||||||||||||||||||||||||||||||||||
Db 296 GGCCTCAAGCATCGACATCGCCGCGGCTTACAAA 333

RESULT 4
B0612189
LOCUS       B0612189 RIKEN full-length enriched, 0 day neonate skin Mus
DEFINITION musculus cDNA clone 4631401M01 5', mRNA sequence.
ACCESSION  B0612189
VERSION     B0612189.1 GI:16453236
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 961)
AUTHORS     Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
             Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
             M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Sasaki
             Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
             D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
             Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
             Muramatsu,M. and Hayashizaki,Y.
             RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
             Unpublished (2001)
TITLE        Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center (GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/
             Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
             M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new

```

genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,K.I., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES
SOURCE
1. .961
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4631401M01"
/clone_lib="RIKEN full-length enriched, 0 day neonate
skin"
/sex="mixed"
/tissue_type="skin"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site.1: Sali; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCACAGCTTTTCTTTTCTTTTCTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTAAATTAATTAATCCGCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I"
BASE COUNT 158 a 302 c 301 g 200 t
ORIGIN
Query Match 1.3%; Score 36; DB 10; Length 961;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 449 ACAGCCAGCGGCGCACACCTTCTCCAGCGCGGGGTG 484
|||||
DB 768 ACAGCCAGCGGCGCACACCTTCTCCAGCGCGGGGTG 803
|||||

RESULT 5
A2842795/c
LOCUS
DEFINITION
2M0141119F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC2M014119 F, DNA sequence.
ACCESSION
A2842795
VERSION
A2842795.1 GI:13012703
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
1 (Bases 1 to 469)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Humil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,I., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0141 row: I column: 19
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 469.
Location/Qualifiers
1. .469
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M014119"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 122 a 151 c 87 g 109 t
ORIGIN
Query Match 1.2%; Score 35; DB 17; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1942 GTCTGTCTCCAGGCAAGTCATCAAGGCTGGCTG 1976
|||||
DB 327 GTCTGTCTCCAGGCAAGTCATCAAGGCTGGCTG 293
|||||

RESULT 6
BE553572
LOCUS
DEFINITION
ur47b04.y1 NCI CGAP Mam2 Mus musculus cDNA clone IMAGE:3153391 5'
similar to TR:054768 054768 A DISINTEGRIN AND METALLOPROTEASE WITH
THROMBOSPONDIN MOTIFS ; mRNA sequence.
ACCESSION
BE553572
VERSION
BE553572.1 GI:9817989
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 522)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-remail.nih.gov

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1056147

Seq primer: -40RP from Gibco
High quality sequence stop: 385.

FEATURES

source Location/Qualifiers
1..522
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:315391"
/clone_lib="NCI_CGAP_Man2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 116 a 133 c 158 g 115 t

ORIGIN
Query Match 1.2%; Score 35; DB 10; Length 522;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1942 GTCTGTCTCCAGGCAAGTGCATCAGGCTGGCTG 1976
|||||
Db 73 GTCTGTCTCCAGGCAAGTGCATCAGGCTGGCTG 107

RESULT 7

BF320986 305 bp mRNA linear EST 29-DEC-2000
LOCUS uz57h10.y1 NCI_CGAP_Man6 Mus musculus CDNA clone IMAGE:3673219 5'
DEFINITION similar to TR:Q9UP80 Q9UP80 METH1 PROTEIN.;, mRNA sequence.

ACCESSION BF320986
VERSION BF320986.1 GI:11270026

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 305)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1433987

High quality sequence stop: 192.
Location/Qualifiers
1..305

FEATURES

source /organism="Mus musculus"
/strain="FVB/N"

/db_xref="taxon:10090"
/clone="IMAGE:3673219"
/clone_lib="NCI_CGAP_Man6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 79 a 71 c 85 g 69 t

ORIGIN

Query Match 1.1%; Score 30; DB 12; Length 305;
Best Local Similarity 100.0%; Pred. NO. 0.005;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1765 GGCTACACACAGCAGCAGCGGCTCACT 1794
|||||

Db 48 GGCTACACACAGCAGCAGCGGCTCACT 77
|||||

RESULT 8

AQ440250 400 bp DNA linear GSS 31-MAR-1999
LOCUS HS_5071.B2.D07.T7A RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate=647 Col=14 Row=H, DNA sequence.

ACCESSION AQ440250
VERSION AQ440250.1 GI:4551589

KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 400)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887

Email: jwallaceu.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (<http://inforesgen.com>). BAC end Web Server:
<http://www.htsc.washington.edu>
Plate: 647 row: H column: 14
Seq primer: T7
Class: BAC ends
High quality sequence stop: 400.

FEATURES

source Location/Qualifiers
1..400
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=647 Col=14 Row=H"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"

/note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"

BASE COUNT 100 a 116 c 98 g 83 t 3 others

ORIGIN

Query Match 1.0%; Score 29; DB 17; Length 400;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2164 GATGACAACACTACCTGGCTCTGAAGACAG 2192

|||||

DB 359 GATGACAACACTACCTGGCTCTGAAGACAG 387

RESULT 9

AW485848

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

PIG.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 415)

FAHRENKRUG, S.C., FRIKING, B.A., ROHRER, G.A., SMITH, T.P.L., CASAS, E.,

STONE, R.T., HEATON, M.P., GROSSE, W.M., BENNETT, G.A., LAEGREID, W.W.

and Keele, J.W.

Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

Unpublished (2000)

Contact: Smith RPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 20

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCAGCTACGAGG

Plate: 28 row: D column: 18

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1..415

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone_lib="MARC LPIG"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;

Library made from pooled tissue from day 11, 13, 15, 20,

and 30 embryos."

BASE COUNT 70 a 141 g 63 t

ORIGIN

Query Match 1.0%; Score 28; DB 10; Length 415;

Best Local Similarity 100.0%; Pred. No. 0.049;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2808 CCGCAAGCCCCAGAGCTGGACTTCTGC 2835

|||||

DB 35 CCGCAAGCCCCAGAGCTGGACTTCTGC 62

RESULT 10

B01986

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

GSS.

B01986

csRL-145A3-u csRL flow sorted Chromosome 11 specific cosmid Homo

sapiens genomic clone csRL-145A3, DNA sequence.

B01986

GI:1411264

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 169)

EVANS, G.A., BURBEE, D., DAVIES, C., HAHNER, L., OLIVER, T., GILBERT, M.,

JONES, D., WARD, T., GILLILAN, E., SCHAGEMANN, J., PROBST, S., HARRIS,

J., DEFORD, J., MCFARLAND, J., BURZINSKI, K., KHAN, M., KUPFER, K. and

Garner, H.R.

Genomic Sequence Sampled Map of Chromosome 11

Unpublished (1996)

Contact: Evans GA, Shane Probst

McDermott Center for Human Growth and Development

University of Texas Southwestern Medical Center At Dallas

5323 Harry Hines Blvd, Dallas TX 75235-8591

Tel: 214-648-1600

Fax: 214-648-1666

Email: gevanseutsw.swmed.edu, shane@mcdermott.swmed.edu

Seq primer: T7

Class: cosmid ends

High quality sequence stop: 169.

Location/Qualifiers

1..169

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="csRL-145A3"

/clone_lib="csRL flow sorted Chromosome 11 specific

cosmid"

/sex="female"

/cell_type="chimeric hamster somatic cell hybrid"

/notes="Vector: sCos-1; Human Chromosome 11 specific cosmid

library prepared from flow sorted human Chromosome 11

derived from Chinese Hamster Ovary (CHO) monochromosomal

somatic cell hybrid, J1"

BASE COUNT 26 a 56 c 41 g 36 t 10 others

ORIGIN

Query Match 0.9%; Score 27; DB 17; Length 169;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CGCGGGGTGTTCGGCGGGCGCTTC 501

|||||

DB 56 CGCGGGGTGTTCGGCGGGCGCTTC 82

BE281680

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 627)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM8537 row: a column: 16

High quality sequence stop: 530.

5

/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 193 a 250 c 250 g 156 t
ORIGIN

Query Match 0.8%; Score 25; DB 13; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1700 TGGAGCCCTGCCCCAGCTCAGCCTC 1724
|||||
Db 540 TGGAGCCCTGCCCCAGCTCAGCCTC 564

RESULT 14
LOCUS BE666088 464 bp mRNA linear EST 25-APR-2001
DEFINITION 149119 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE666088
VERSION BE666088.1 GI:10026679
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 464)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCCGAGTCACGACG
Plate: 62 row: B column: 2
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1. .464
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from day 20 and day 40 embryos."

BASE COUNT 73 a 157 c 163 g 71 t
ORIGIN

Query Match 0.8%; Score 24; DB 10; Length 464;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2812 AAGCCCCAGGAGCTGGACTTCTGC 2835
|||||
Db 215 AAGCCCCAGGAGCTGGACTTCTGC 238

RESULT 15
LOCUS BI692986 664 bp mRNA linear EST 18-SEP-2001
DEFINITION 603344521F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5372261 5', mRNA sequence.
ACCESSION BI692986
VERSION BI692986.1 GI:15655615
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 664)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1947 row: j column: 06
High quality sequence stop: 664.

Location/Qualifiers
1. .664
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 133 a 168 c 207 g 156 t
ORIGIN

Query Match 0.8%; Score 24; DB 13; Length 664;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2752 CTCAGTGTGTGGCCACGAGGC 2775
|||||
Db 46 CTCAGTGTGTGGCCACGAGGC 69

Search completed: May 16, 2003, 04:02:41
Job time : 3610 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 22:53:01 ; Search time 340 Seconds
(without alignments)
10824.099 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgcttctgctggcctcct.....gcgtctgagccgtgtgta 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NHW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	2853	100.0	2853	10	US-09-965-631-3
2	2853	100.0	3446	10	US-09-965-631-7
3	2298	80.5	2469	9	US-10-163-316-3
4	2298	80.5	2940	9	US-10-163-316-1
5	1091.2	38.2	1104	10	US-09-965-631-5
6	959	33.6	966	10	US-09-965-631-1
7	634.4	22.2	2184	9	US-10-097-597-13
8	634.4	22.2	2184	9	US-10-097-580-13
9	634.4	22.2	2184	10	US-09-445-023A-13
10	627.2	22.0	2184	9	US-10-097-597-2
11	627.2	22.0	2184	9	US-10-097-580-2
12	627.2	22.0	2184	10	US-09-445-023A-2
13	626.6	22.0	4676	12	US-10-105-929-1
14	590.4	20.7	3638	10	US-09-918-171A-8
15	581.6	20.4	4407	9	US-10-174-590-351
16	581.6	20.4	4407	9	US-10-176-758-351
17	581.6	20.4	4407	9	US-10-175-737-351
18	581.6	20.4	4407	9	US-10-173-706-351
19	581.6	20.4	4407	9	US-10-175-738-351

20	581.6	20.4	4407	9	US-10-175-752-351	Sequence 351, App
21	581.6	20.4	4407	9	US-10-176-482-351	Sequence 351, App
22	581.6	20.4	4407	9	US-10-176-757-351	Sequence 351, App
23	581.6	20.4	4407	9	US-10-176-913-351	Sequence 351, App
24	581.6	20.4	4407	9	US-10-180-552-351	Sequence 351, App
25	581.6	20.4	4407	9	US-10-180-557-351	Sequence 351, App
26	581.6	20.4	4407	9	US-10-173-700-351	Sequence 351, App
27	581.6	20.4	4407	9	US-10-174-572-351	Sequence 351, App
28	581.6	20.4	4407	9	US-10-174-579-351	Sequence 351, App
29	581.6	20.4	4407	9	US-10-174-582-351	Sequence 351, App
30	581.6	20.4	4407	9	US-10-174-588-351	Sequence 351, App
31	581.6	20.4	4407	9	US-10-175-739-351	Sequence 351, App
32	581.6	20.4	4407	9	US-10-175-740-351	Sequence 351, App
33	581.6	20.4	4407	9	US-10-175-743-351	Sequence 351, App
34	581.6	20.4	4407	9	US-10-176-488-351	Sequence 351, App
35	581.6	20.4	4407	9	US-10-176-492-351	Sequence 351, App
36	581.6	20.4	4407	9	US-10-176-747-351	Sequence 351, App
37	581.6	20.4	4407	9	US-10-176-750-351	Sequence 351, App
38	581.6	20.4	4407	9	US-10-176-985-351	Sequence 351, App
39	581.6	20.4	4407	9	US-10-176-987-351	Sequence 351, App
40	581.6	20.4	4407	9	US-10-176-991-351	Sequence 351, App
41	581.6	20.4	4407	9	US-10-176-992-351	Sequence 351, App
42	581.6	20.4	4407	9	US-10-176-993-351	Sequence 351, App
43	581.6	20.4	4407	9	US-10-184-658-351	Sequence 351, App
44	581.6	20.4	4407	9	US-10-173-695-351	Sequence 351, App
45	581.6	20.4	4407	9	US-10-173-697-351	Sequence 351, App

ALIGNMENTS

RESULT 1

US-09-965-631-3
; Sequence 3, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-3

Query Match	100.0%;	Score 2853;	DB 10;	Length 2853;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2853;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCTTCTGCTGGGCATCCTAAACCTGGCTTCCCGGGCGGAACCGCTGGAGCTT	1	JAG 60
Db	1	ATGCTTCTGCTGGGCATCCTAAACCTGGCTTCCCGGGCGGAACCGCTGGAGCTT	1	JAG 60
QY	61	CCAGAGCGGGAGGTAGTCGTTCCATCCGACTGGACCCGACATTAACGGCCGCCATAC	120	
Db	61	CCAGAGCGGGAGGTAGTCGTTCCATCCGACTGGACCCGACATTAACGGCCGCCATAC	120	
QY	121	TACTGGCGGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTTCAGATCACAGCATTT	180	
Db	121	TACTGGCGGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTTCAGATCACAGCATTT	180	
QY	181	CAGAGGAGCTTTTACCTACACCTGACCGCGGATGCTCAGTTTGGTCCCGCTTCTCC	240	
Db	181	CAGAGGAGCTTTTACCTACACCTGACCGCGGATGCTCAGTTTGGTCCCGCTTCTCC	240	
QY	241	ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGGCTTCTCAGACCTTCGACGC	300	

|||||
Db 241 ACTGACATCTGGCGTCCCTCCAGGGCTCACCGGGGTCTTCAGACTGGAGCC 300
QY 301 TGCTTCTATTCTGGGACGTGAACGCCGAGCGGACTCGTTCGCTGCTGTGAGCCCTGTGC 360
Db 301 TGCTTCTATTCTGGGACGTGAACGCCGAGCGGACTCGTTCGCTGCTGTGAGCCCTGTGC 360
QY 361 GGGGGCTCCGGGAGCCCTTGGCTACCGAGGCGCGGATATGTATAGCCCGCTGCC 420
Db 361 GGGGGCTCCGGGAGCCCTTGGCTACCGAGGCGCGGATATGTATAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGGAGCAACAGCAACAGCGGCGCACACTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGCGGAGCAACAGCAACAGCGGCGCACACTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCTCGTGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCTCGTGGGGTGGCTCGGGCTGG 540
QY 541 AACCCGCCATCTACGGGCTTGAGCCCTTACAGCCCGCGGCGGGCTTCGGGGAG 600
Db 541 AACCCGCCATCTACGGGCTTGAGCCCTTACAGCCCGCGGCGGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGGCGAGGCTTGGCGCGCAAGCGCTTTCGCTCTATCCCGGCTACGTG 660
Db 601 AGTCGTAGCGCGGCGAGGCTTGGCGCGCAAGCGCTTTCGCTCTATCCCGGCTACGTG 660
QY 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGGTCAAGTTCACCGGCGGACCTGGAACAT 720
Db 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGGTCAAGTTCACCGGCGGACCTGGAACAT 720
QY 721 TATCTGCTGACGCTGTGGACAGCGGCGGCGACTACCGGCATCCAGCATCTCAAC 780
Db 721 TATCTGCTGACGCTGTGGACAGCGGCGGCGACTACCGGCATCCAGCATCTCAAC 780
QY 781 CCCATCAACATCTGTGTGTCAGGTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
Db 781 CCCATCAACATCTGTGTGTCAGGTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
QY 841 GTCACCGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 900
Db 841 GTCACCGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 900
QY 901 AAGTGAAGTACAGACACCCGAGTCTGGGACCTGCCATCTCTTACAGGACGAG 960
Db 901 AAGTGAAGTACAGACACCCGAGTCTGGGACCTGCCATCTCTTACAGGACGAG 960
QY 961 CTGTGTGGAGCCACCACTGTGACACCTTGGCGATGGCTGTGTGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGACACCTTGGCGATGGCTGTGTGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGTGTCTGTCAATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1080
Db 1021 CCCAAGAGAAGTGTCTGTCAATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1080
QY 1081 CAGGACTGGCCACGTGTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTG 1140
Db 1081 CAGGACTGGCCACGTGTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCTCATCCAGATCCAGCTGCC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCTCATCCAGATCCAGCTGCC 1200
QY 1201 AACCCCTGTGAGCTGTGCTGTCAATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1260
Db 1201 AACCCCTGTGAGCTGTGCTGTCAATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1260
QY 1261 GACTGCTCTGAGCAACCCAGGACCCATCTCCCTCCCGAGGATCTGCCGGCGCC 1320
Db 1261 GACTGCTCTGAGCAACCCAGGACCCATCTCCCTCCCGAGGATCTGCCGGCGCC 1320
QY 1321 AGCTACACCTGTAGCAGCAGTGCAGCTGGCTTTTGGCGTGGCTTCCAGCCCTGTGCT 1380
|||||

Db 1321 AGCTACACCTGTAGCAGCAGTGCAGCTGGCTTTTGGCGTGGCTCCAAGCCCTGTCTCT 1380
QY 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1440
Db 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1440
QY 1441 CAGACCCGCCACTTCCCTTGGCGGATGGCACCAAGCTGTGGCGAGGCAAGCTCTGCCTC 1500
Db 1441 CAGACCCGCCACTTCCCTTGGCGGATGGCACCAAGCTGTGGCGAGGCAAGCTCTGCCTC 1500
QY 1501 AAAGGGGCTGCGTGGAGAGACACAACCTCAACAAGCACAGGCTGGATGCTCTCGGCC 1560
Db 1501 AAAGGGGCTGCGTGGAGAGACACAACCTCAACAAGCACAGGCTGGATGCTCTCGGCC 1560
QY 1561 AAATGGGATCCCTATGCCCCCTGCTCGCGCACATGTGGTGGGGCGTGCAGCTGGT*AGG 1620
Db 1561 AAATGGGATCCCTATGCCCCCTGCTCGCGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1620
QY 1621 AGGCAGTGCACCAACCCACCCCTGCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
Db 1621 AGGCAGTGCACCAACCCACCCCTGCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCTGCCCCAGCTCAGCCTCCGGAAAGAGTTCGGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCTGCCCCAGCTCAGCCTCCGGAAAGAGTTCGGG 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTGCC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTGCC 1800
QY 1801 GTGCATGGGTGCCAAGTACTCGGGCGTGTCTCCCGGACAAAGTCAAGCTCACTGTGC 1860
Db 1801 GTGCATGGGTGCCAAGTACTCGGGCGTGTCTCCCGGACAAAGTCAAGCTCACTGTGC 1860
QY 1861 CGAGCCAAATGCACTGCTTCTATGTGCTGCAACCCCAAGTGGTGACGCGCACCTG 1920
Db 1861 CGAGCCAAATGCACTGCTTCTATGTGCTGGCACCCCAAGTGGTGAGCGGACGCTG 1920
QY 1921 TGCTCTCTGACTCCACTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGGAACTGGGCTCCAGAAGAGATTGCAAGTGTGGGTGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGGCTCCAGAAGAGATTGCAAGTGTGGGTGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGTGTGACTTTCACCAAGCCCATGATGCTGCTACAAATTTTCGTTGGT 2100
Db 2041 AGCTGCAAGAAGTGTGACTTTCACCAAGCCCATGATGCTGCTACAAATTTTCGTTGGT 2100
QY 2101 GCCATCCCGGAGGCGCTCAAGCATCGCATCGCCAGCGCGGTTCACAAAGGCTGATC 2160
Db 2101 GCCATCCCGGAGGCGCTCAAGCATCGCATCGCCAGCGCGGTTCACAAAGGCTGATC 2160
QY 2161 GGGATGACAACTACCTGGCTCTGAAGAAGCAAGCAAGTACCTGCTCAAGGGCAT 2220
Db 2161 GGGATGACAACTACCTGGCTCTGAAGAAGCAAGCAAGTACCTGCTCAAGGGCAT 2220
QY 2221 TTCTGTGTGTGGCGGTGGAGCGGACCTGGTGTGAAGGGCAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCTGTGTGTGGCGGTGGAGCGGACCTGGTGTGAAGGGCAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCACGGGCACAGCGGTGGAGAGCTCGAGCTTCCGGGCCCATCTCGAGCGCGTGACC 2340
Db 2281 GGCACGGGCACAGCGGTGGAGAGCTCGAGCTTCCGGGCCCATCTCGAGCGCGTGACC 2340
QY 2341 GTGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGTCCGCTACTCTCTTCTATCTG 2400
Db 2341 GTGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGTCCGCTACTCTCTTCTATCTG 2400
QY 2401 CCCAAAGAGCCTCGGGAGGACAAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
Db 2401 CCCAAAGAGCCTCGGGAGGACAAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTGTC 2460

QY 2461 TTGCACACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCAGCGCGAGCAGCCCCCT 2520
|||||
Db 2461 TTGCACACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCAGCGCGAGCAGCCCCCT 2520
|||||
QY 2521 GCACCTGGGTGGGTGGAGCTGGGGCCGTGCTCCGGGAGCTGGCGAGTGGCTGCAG 2580
|||||
Db 2521 GCACCTGGGTGGGTGGAGCTGGGGCCGTGCTCCGGGAGCTGGCGAGTGGCTGCAG 2580
|||||
QY 2581 AAGCGGGCGGTGGAGCTGTCGGGGCTCCCGCGGAGCGACGTCCTGCTGCTGATGCA 2640
|||||
Db 2581 AAGCGGGCGGTGGAGCTGTCGGGGCTCCCGCGGAGCGACGTCCTGCTGCTGATGCA 2640
|||||
QY 2641 GCCATCGCCCGGTGGAGACACAAAGCTGCGGGAGCGCTCCCGACCTGGAGCTCAGC 2700
|||||
Db 2641 GCCATCGCCCGGTGGAGACACAAAGCTGCGGGAGCGCTCCCGACCTGGAGCTCAGC 2700
|||||
QY 2701 GCCTGGTCACCTGCTCCAAAGCTGCGGGCGGGGATTCAGAGCGGCTCACTCAAGTGT 2760
|||||
Db 2701 GCCTGGTCACCTGCTCCAAAGCTGCGGGCGGGGATTCAGAGCGGCTCACTCAAGTGT 2760
|||||
QY 2761 GTGGGCCACGAGCGCGCTGCTGGCCCGGACCAAGTGCACCGCAAGCCCGCAG 2820
|||||
Db 2761 GTGGGCCACGAGCGCGCTGCTGGCCCGGACCAAGTGCACCGCAAGCCCGCAG 2820
|||||
QY 2821 GAGCTGGAGTCTGCTGCTCTGAGGCGGCTGCTGA 2853
|||||
Db 2821 GAGCTGGAGTCTGCTGCTCTGAGGCGGCTGCTGA 2853
|||||

RESULT 2

US-09-965-631-7
; Sequence 7, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; FILE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3446
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-7

Query Match 100.0%; Score 2853; DB 10; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTCTGGGATCCTAACCTTGGCTTTGCGCGGCGAACCGCTGGAGGCTCTGAG 60
|||||
Db 397 ATGCTTCTCTGGGATCCTAACCTTGGCTTTGCGCGGCGAACCGCTGGAGGCTCTGAG 456
|||||
QY 61 CCAGACGGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGGCGCGCGCTAC 120
|||||
Db 457 CCAGACGGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGGCGCGCGCTAC 516
|||||
QY 121 TACTGCGGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCACACATTT 180
|||||
Db 517 TACTGCGGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCACACATTT 576
|||||
QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
|||||
Db 577 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 636
|||||
QY 241 ACTGAGCATCTGGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGGAGCG 300
|||||

Db 637 ACTGAGCATCTGGGGCTCCCGCTCCAGGGGCTCACCGGGGGCTCTTCACACCTGCAACGC 696
|||||
QY 301 TGCTTCTATTTCTGGGACGTGAACGCCGAGCCGAGTCTGCTGCTGTGAGCCCTGTGC 360
|||||
Db 697 TGCTTCTATTTCTGGGACGTGAACGCCGAGCCGAGTCTGCTGCTGTGAGCCCTGTGC 756
|||||
QY 361 GGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCGCTTGCCC 420
|||||
Db 757 GGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCGCTTGCCC 816
|||||
QY 421 AATGCTAGCCGCCGCCGGCGAGCGACACAGCAGGGGSCACACCTTCTCCAGCCCGG 480
|||||
Db 817 AATGCTAGCCGCCGCCGGCGAGCGACACAGCCAGGGGSCACACCTTCTCCAGCCCGG 876
|||||
QY 481 GGTGTTCCCGGGCGGGCTTCCGGAGACCCACACCTCTCGCTGCGGGGTGSCCTCGGCTG 540
|||||
Db 877 GGTGTTCCCGGGCGGGCTTCCGGAGACCCACACCTCTCGCTGCGGGGTGSCCTCGGCTG 936
|||||
QY 541 AACCCGCCATCTACGGGCCCTGGACCCCTTACAAGCCGGCGGGGGCTTCGCGGAG 600
|||||
Db 937 AACCCGCCATCTACGGGCCCTGGACCCCTTACAAGCCGGCGGGGGCTTCGCGGAG 996
|||||
QY 601 AGTCGTAGCCGCCGAGTCTGGCGCGCCAAAGCTTTCTGCTATCCCGCGGTACGTG 660
|||||
Db 997 AGTCGTAGCCGCCGAGTCTGGCGCGCCAAAGCTTTCTGCTATCCCGCGGTACGTG 1056
|||||
QY 661 GAGACGCTGGTGGTTCGCGGACGAGTCAATGGTCAAGTTCACCGCGCGGACCTTGGAAACAT 720
|||||
Db 1057 GAGACGCTGGTGGTTCGCGGACGAGTCAATGGTCAAGTTCACCGCGCGGACCTTGGAAACAT 1116
|||||
QY 721 TATCTGCTGAGCTGTGGCAACGGCGCGGACTCTACCGCCATCCACGATCTCTCAAC 780
|||||
Db 1117 TATCTGCTGAGCTGTGGCAACGGCGCGGACTCTACCGCCATCCACGATCTCTCAAC 1176
|||||
QY 781 CCATCAACATCGTTGGTCAAGTCTGCTCTTATAGATCGTGACTCCGGCGCCCAAG 840
|||||
Db 1177 CCATCAACATCGTTGGTCAAGTCTGCTCTTATAGATCGTGACTCCGGCGCCCAAG 1236
|||||
QY 841 GTCACCGCAATCGGGCCCTGACGCTCGCAACTTCTGTCCTGGGAGAGAAAGCTGAAC 900
|||||
Db 1237 GTCACCGCAATCGGGCCCTGACGCTCGCAACTTCTGTCCTGGGAGAGAAAGCTGAAC 1296
|||||
QY 901 AAGTGAGTGACAGCACACCCGAGTACTGGGACACTCCCATCTCTTCCACGACGAGAC 960
|||||
Db 1297 AAGTGAGTGACAGCACACCCGAGTACTGGGACACTCCCATCTCTTCCACGACGAGAC 1356
|||||
QY 961 CTGTGTGAGCCACCACTGTGACACCTGCGGCACTGCTGATGGGTACCATGTGTGAC 1020
|||||
Db 1357 CTGTGTGAGCCACCACTGTGACACCTGCGGCACTGCTGATGGGTACCATGTGTGAC 1416
|||||
QY 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGACGATGGGTTCCATCAGCCTTCCACATTCGCC 1080
|||||
Db 1417 CCCAAGAGAGCTGCTCTGCTATTGAGGACGATGGGTTCCATCAGCCTTCCACATTCGCC 1476
|||||
QY 1081 CACGAGCTGGCCACGCTGTTCAACATGCCCATCACAATGTGAAAGTCTGTGAGGAGGTG 1140
|||||
Db 1477 CACGAGCTGGCCACGCTGTTCAACATGCCCATCACAATGTGAAAGTCTGTGAGGAGGTG 1536
|||||
QY 1141 TTTGGAGAGCTCCGAGCGCAACACATGATGCCCGGACCCCTCATCCAGATCGACCGTGC 1200
|||||
Db 1537 TTTGGAGAGCTCCGAGCGCAACACATGATGCCCGGACCCCTCATCCAGATCGACCGTGC 1596
|||||
QY 1201 AACCCCTGGTCCAGCTGCGAGTGTGCCATCATCACCAGCTTCTGAGAGCGGCGACGGT 1260
|||||
Db 1597 AACCCCTGGTCCAGCTGCGAGTGTGCCATCATCACCAGCTTCTGAGAGCGGCGACGGT 1656
|||||
QY 1261 GACTGCTCTTGACCAACCCAGCAAGCCATCTCCCTGCCCGGAGGATCTGCCGGCGCC 1320
|||||
Db 1657 GACTGCTCTTGACCAACCCAGCAAGCCATCTCCCTGCCCGGAGGATCTGCCGGCGCC 1716
|||||
QY 1321 AGCTACACCTTGACCGAGCTGCGAGTGGCTTTTGGCGGTGGGCTCCAAGCCCTGTCT 1380
|||||
Db 1717 AGCTACACCTTGACCGAGCTGCGAGTGGCTTTTGGCGGTGGGCTCCAAGCCCTGTCT 1776
|||||

```
QY 1381 TACATGCAGTACTGCACCAAGCTGTGGTGCACCGGGAGGCCAAGGACACAGATGGTGTGC 1440
Db 1777 TACATGCAGTACTGCACCAAGCTGTGGTGCACCGGGAGGCCAAGGACACAGATGGTGTGC 1836
QY 1441 CAGACCCGCCACTTCCCTCGTGGCCGATGGACACAGCTGTGGGAGGCCAAGCTGTGCCTC 1500
Db 1837 CAGACCCGCCACTTCCCTCGTGGCCGATGGACACAGCTGTGGGAGGCCAAGCTGTGCCTC 1896
QY 1501 AAAGGGGCTCGTGGAGAGACAAACCTCAACAAAGCACAGGCTGGATGCTTCTTGGGCC 1560
Db 1897 AAAGGGGCTCGTGGAGAGACAAACCTCAACAAAGCACAGGCTGGATGCTTCTTGGGCC 1956
QY 1561 AAATGGGATCCTATGATGCGCCCTGTCCGACACATGTGGTGGGCGGTGACAGTGGCCAGG 1620
Db 1957 AAATGGGATCCTATGATGCGCCCTGTCCGACACATGTGGTGGGCGGTGACAGTGGCCAGG 2016
QY 1621 AGGCAGTGCACCAACCCCTGCGCAACGGGGCAAGTACTCGGAGGGAGTGAAGGGTG 1680
Db 2017 AGGCAGTGCACCAACCCCTGCGCAACGGGGCAAGTACTCGGAGGGAGTGAAGGGTG 2076
QY 1681 AAATACCGATCCTGCAATCTGGAGCCTGCCCGAGCTCAGCCTCCGGAAGAGCTTCCGG 1740
Db 2077 AAATACCGATCCTGCAATCTGGAGCCTGCCCGAGCTCAGCCTCCGGAAGAGCTTCCGG 2136
QY 1741 GAGGAGCAGTGTAGGGCTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTCGCC 1800
Db 2137 GAGGAGCAGTGTAGGGCTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTCGCC 2196
QY 1801 GTGCAATGGGTGCCAAAGTACTCCGGGCTGTCTCCCGGGACAAAGTGCAGGCTCACTCTGC 1860
Db 2197 GTGCAATGGGTGCCAAAGTACTCCGGGCTGTCTCCCGGGACAAAGTGCAGGCTCACTCTGC 2256
QY 1861 CGAGCCAAATGGCACTGCTACTTCTATGTGCTGGCACCACCAAGTGTGGACGCGACGCTG 1920
Db 2257 CGAGCCAAATGGCACTGCTACTTCTATGTGCTGGCACCACCAAGTGTGGACGCGACGCTG 2316
QY 1921 TGCTCTCTGACTCCACTCCGCTGTGTCACAGGCAAGTGCATTAAGCTGGCTGTGAT 1980
Db 2317 TGCTCTCTGACTCCACTCCGCTGTGTCACAGGCAAGTGCATTAAGCTGGCTGTGAT 2376
QY 1981 GGGAACTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAAATAAG 2040
Db 2377 GGGAACTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAAATAAG 2436
QY 2041 AGCTGCAAGAGTGTGACTGACTTTCACCAAGCCCATGATGGCTACAAATTTCTGGTGTG 2100
Db 2437 AGCTGCAAGAGTGTGACTGACTTTCACCAAGCCCATGATGGCTACAAATTTCTGGTGTG 2496
QY 2101 GCCATCCCCGAGCGCCTCAAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2160
Db 2497 GCCATCCCCGAGCGCCTCAAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2556
QY 2161 GGGGATGACAACCTACCTGGCTCTGAAGAAGACCAAGGCAAGTACCTGCTCAACGGGAT 2220
Db 2557 GGGGATGACAACCTACCTGGCTCTGAAGAAGACCAAGGCAAGTACCTGCTCAACGGGAT 2616
QY 2221 TTGCTGTGTGCGGCTGGAGCGGGACCTGGTGGTGAAGGGCACTGCTGCGGTACAGC 2280
Db 2617 TTGCTGTGTGCGGCTGGAGCGGGACCTGGTGGTGAAGGGCACTGCTGCGGTACAGC 2676
QY 2281 GGCACGGGCACAGCGTGGAGAGCCCTGACAGGCTTCCCGGCCCATCTCGAGCGCGTGAAC 2340
Db 2677 GGCACGGGCACAGCGTGGAGAGCCCTGACAGGCTTCCCGGCCCATCTCGAGCGCGTGAAC 2736
QY 2341 GTGAGGTCTCTCTGGTGGGAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2737 GTGAGGTCTCTCTGGTGGGAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAAGAGCCTGGGAGGACAGTCTCTCATCCAGGACCCCGGGGACCCCTCTGTC 2460
Db 2797 CCCAAAGAGCCTGGGAGGACAGTCTCTCTCATCCAGGACCCCGGGGACCCCTCTGTC 2856
```

```
QY 2461 TTGCACAACAGCGTCTCTAGCCCTCTCCAAACAGGTGGAGCAGCCGAGACAGAGCCGCCCT 2520
Db 2857 TTGCACAACAGCGTCTCTAGCCCTCTCCAAACAGGTGGAGCAGCCGAGACAGAGCCGCCCT 2916
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGGCGGTGCTCCGGAGAGTGGGAGTGGCCCTGCAG 2580
Db 2917 GCACGCTGGGTGGCTGGCAGCTGGGGGCGGTGCTCCGGAGAGTGGGAGTGGCCCTGCAG 2976
QY 2581 AAGCGGGCGGTGACTGTCTGGGGCTCCCGCGGACAGCCAGGTCCTTGCCTGTGATGCA 2640
Db 2977 AAGCGGGCGGTGACTGTCTGGGGCTCCCGCGGACAGCCAGGTCCTTGCCTGTGATGCA 3036
QY 2641 GCCATCGGCCCGTGGAGACAAAGCTTGGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
Db 3037 GCCATCGGCCCGTGGAGACAAAGCTTGGGGGAGCCCTGCCACCTGGGAGCTCAGC 3096
QY 2701 GCCTGGTCACTTCTTCAAGAGCTGGGGCGGGGATTTTCAGAGGCGCTCACTCAAGTGT 2760
Db 3097 GCCTGGTCACTTCTTCAAGAGCTGGGGCGGGGATTTTCAGAGGCGCTCACTCAAGTGT 3156
QY 2761 GTGGGCCACGGAGCGCGTGTGGCCCGGAGCACAGTGCACCTTGCACCGCAAGCCCGAG 2820
Db 3157 GTGGGCCACGGAGCGCGTGTGGCCCGGAGCACAGTGCACCTTGCACCGCAAGCCCGAG 3216
QY 2821 GAGCTGGACTTCTTCTGCTCTGAGGCGGTGCTGA 2853
Db 3217 GAGCTGGACTTCTTCTGCTCTGAGGCGGTGCTGA 3249
```

RESULT 3

```
US-10-163-316-3
; Sequence 3, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-10-163-316-3
```

Query Match

```
Best Local Similarity 80.5%; Score 2298; DB 9; Length 2469;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;
```

```
QY 1 ATGCTTCTGTGGCATCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGTGGCATCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGGGGAGGTAGTCTCCATCCGACTGGAGCCCGGACATTAACGGCCGCGGTAC 120
Db 61 CCAGAGGGGAGGTAGTCTCCATCCGACTGGAGCCCGGACATTAACGGCCGCGGTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT 180
Db 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGACCGGATGCTCAGTCTTGGCTCCCGGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACCGGATGCTCAGTCTTGGCTCCCGGCTTCTCC 240
```

QY 241 ACTGAGCATCGGCGCTCCCTCCAGGGCTCACGGGCTCCTCAGACCTCGGACGC 300
DB 241 ACTGAGCATCGGCGCTCCCTCCAGGGCTCACGGGCTCCTCAGACCTCGGACGC 300
QY 301 TGCCTTCTATTCTGGGACGTGAACCGCGAGCCGACCTCGTTCGCTGTGTGAGCCTGTGC 360
DB 301 TGCCTTCTATTCTGGGCGGTGAACCGCGAGCCGACCTCGTTCGCTGTGTGAGCCTGTGC 360
QY 361 GGGGGCTCCCGGAGCCTTTGGCTACGAGGCGCCGAGTATGTCATTAGCCGCTGCC 420
DB 361 GGGGGCTCCCGGAGCCTTTGGCTACGAGGCGCCGAGTATGTCATTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGGCGCAGCGCAACAGCGAGCGGCACACCTTCTCCAGCGCGG 480
DB 421 AATGCTAGCGCGCGCGGCGCAGCGCAACAGCGAGCGGCACACCTTCTCCAGCGCGG 480
QY 481 GGTGTCCGGCGGGGCTTCGGAGACCCACCTCTCGCTGCGGGGTGGCTCGGGGTGG 540
DB 481 GGTGTCCGGCGGGGCTTCGGAGACCCACCTCTCGCTGCGGGGTGGCTCGGGGTGG 540
QY 541 AACCCCGCATCTTACGGGCGCTGGACCTTTACAAGCGCGCGGGCGGCTTCGGGGAG 600
DB 541 AACCCCGCATCTTACGGGCGCTGGACCTTTACAAGCGCGCGGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCAGCTCTGGCGCGCCCAAGCTTTCTGTCTATCCCGCGGTACGTG 660
DB 601 AGTCGTAGCGCGCAGCTCTGGCGCGCCCAAGCTTTCTGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGGTGTGGCGGACGAGTCATAGTTCAGTTCACGGCGCGGACCTTGGACAT 720
DB 661 GAGACGCTGGTGTGGCGGACGAGTCATAGTTCAGTTCACGGCGCGGACCTTGGACAT 720
QY 721 TATCTGTGACGCTGCTGGCAACGGCGCGGACTCTACCGGCAATCCAGCATCTCAAC 780
DB 721 TATCTGTGACGCTGCTGGCAACGGCGCGGACTCTACCGGCAATCCAGCATCTCAAC 780
QY 781 CCATCAACATCTGTGTGGTCAAGTGTCTTCTTAGAGATCTGTGACTCCGGGCCAAG 840
DB 781 CCATCAACATCTGTGTGGTCAAGTGTCTTCTTAGAGATCTGTGACTCCGGGCCAAG 840
QY 841 GTACCGGCAATCGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAAGAACTGAAC 900
DB 841 GTACCGGCAATCGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAAGAACTGAAC 900
QY 901 AAAGTGAGTGACAAACCCGAGTACTGGACACTGCCATCTTCCACGACGAGGAC 960
DB 901 AAAGTGAGTGACAAACCCGAGTACTGGACACTGCCATCTTCCACGACGAGGAC 960
QY 961 CTGTGTGGAGCCACCACTGTGACACCTTGGCATGGCTGATGTGGGTACCATGTGTGAC 1020
DB 961 CTGTGTGGAGCCACCACTGTGACACCTTGGCATGGCTGATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGTGTCTGTCTATTAGAGACGATGGGCTTCCATCAGCCCTTCCACCTGCC 1080
DB 1021 CCCAAGAGAAGTGTCTGTCTATTAGAGACGATGGGCTTCCATCAGCCCTTCCACCTGCC 1080
QY 1081 CAGGAGCTGGGCGAGCTGTTCACATGCCCATGCCATGACATCTGAAGTCTGTGAGGAGTG 1140
DB 1081 CAGGAGCTGGGCGAGCTGTTCACATGCCCATGCCATGACATCTGAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACGCTGCC 1200
DB 1141 TTTGGGAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACGCTGCC 1200
QY 1201 AACCCCTGGTCAAGCTGCACTGCCATCATCAGCCAGTTCCTGGACAGCGGACCGGT 1260
DB 1201 AACCCCTGGTCAAGCTGCACTGCCATCATCAGCCAGTTCCTGGACAGCGGACCGGT 1260
QY 1261 GACTGCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTCCCGGGGCC 1320
DB 1261 GACTGCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTCCCGGGGCC 1320
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380

DB 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCT 1380
QY 1381 TACATGCACTACTGCAACCAAGCTGTGGTGCACCCGGGAGGCCAAGGACAGATGTGTGC 1440
DB 1381 TACATGCACTACTGCAACCAAGCTGTGGTGCACCCGGGAGGCCAAGGACAGATGTGTGC 1440
QY 1441 CAGACCCGCACTTCCCTGGCGGATGACACAGCTGTGGGAGGGCAAGCTTCCCTC 1500
DB 1441 CAGACCCGCACTTCCCTGGCGGATGACACAGCTGTGGGAGGGCAAGCTTCCCTC 1500
QY 1501 AAAGGGCCTGCTGGAGAGACACACCTTCAACAAAGCACAG----- 1541
DB 1501 AAAGGGCCTGCTGGAGAGACACACCTTCAACAAAGCACAGCTCCGACTGACATCAT 1560
QY 1542 -----GGTGGATGT 1551
DB 1561 TCTCCAAAACAACTCTTTAAGGCTTACCAATGGCTGCACACTACACAGGTGGATGT 1620
QY 1552 TCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCGCACATGTGTGGGGCCGTGCAG 1611
DB 1621 TCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCGCACATGTGTGGGGCCGTGCAG 1680
QY 1612 CTGCCAGGAGGAGTGCACCAACCCCTGCCAACCGGGGGCAAGTACTGCGAGGGA 1671
DB 1681 CTGCCAGGAGGAGTGCACCAACCCCTGCCAACCGGGGGCAAGTACTGCGAGGGA 1740
QY 1672 GTGAGGTGAATACCGATCCTCAATCTGGAGCCTGCCAGCTCAGCCTCCGGAAAG 1731
DB 1741 GTGAGGTGAATACCGATCCTCAACCTGGAGCCTGCCAGCTCAGCCTCCGGAAAG 1800
QY 1732 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTACAACACACAGCAACACCGGCTC 1791
DB 1801 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTACAACACACAGCAACACCGGCTC 1860
QY 1792 ACTCTCGCCCTGGCATGGGTGCCAAAGTACTCCGGCGTGTCTCCCGGGACAAAGTCAAG 1851
DB 1861 ACTCTCGCCCTGGCATGGGTGCCAAAGTACTCCGGCGTGTCTCCCGGGACAAAGTCAAG 1920
QY 1852 CTATCTCGCGAGGCAATGGCACTGGCTACTTATGTGTGCTGCACCCAGAGTGGTGGAC 1911
DB 1921 CTATCTCGCGAGGCAATGGCACTGGCTACTTATGTGTGCTGCACCCAGAGTGGTGGAC 1980
QY 1912 GGACAGCTGTGCTCTCTGACTCCACTCCGCTGTGTCCAAAGGCAAGTGCATCAAGGCT 1971
DB 1981 GGACAGCTGTGCTCTCTGACTCCACTCCGCTGTGTCCAAAGGCAAGTGCATCAAGGCT 2040
QY 1972 GGCTGTGATGGGAACCTGGGCTCCAAAGAGAGATTGACAAAGTGTGGGTGTGGGGGA 2031
DB 2041 GGCTGTGATGGGAACCTGGGCTCCAAAGAGAGATTGACAAAGTGTGGGTGTGGGGGA 2100
QY 2032 GACATTAAGAGCTGCAAGAAGTGAAGTGGCTTCCACCAAGCCCATGCATGCTACAAT 2091
DB 2101 GACATTAAGAGCTGCAAGAAGTGAAGTGGCTTCCACCAAGCCCATGCATGCTACAAT 2160
QY 2092 TTCGTGTGGCCATCCCGCGAGCGCTCAAGCATCGACATCCCGCAGCGGTTACAA 2151
DB 2161 TTCGTGTGGCCATCCCGCGAGCGCTCAAGCATCGACATCCCGCAGCGGTTACAA 2220
QY 2152 GGGCTGATCGGGGATGACAACTACCTGGCTTGAAGAACAGCCAAAGTACCTTGCTC 2211
DB 2221 GGGCTGATCGGGGATGACAACTACCTGGCTTGAAGAACAGCCAAAGTACCTTGCTC 2280
QY 2212 AACGGGATTTCTGTGTGTCGGGGTGGAGCGGACCTGGTGTGAGGCACTGTGCTG 2271
DB 2281 AACGGGATTTCTGTGTGTCGGGGTGGAGCGGACCTGGTGTGAGGCACTGTGCTG 2340
QY 2272 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCTGCAGGCTTCCCGGCCCATCTCGGAG 2331
DB 2341 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCTGCAGGCTTCCCGGCCCATCTCGGAG 2400
QY 2332 CCGGTGACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGC 2385
DB 2400 CCGGTGACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGC 2385

Db 2401 CGCGTACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGGCCCGGACCTGCG 2454

RESULT 4
US-10-163-316-1
; Sequence 1, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MPI01-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIORITY FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIORITY FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940)
; OTHER INFORMATION: n = A,T,C or G
US-10-163-316-1

Query Match 80.5%; Score 2298; DB 9; Length 2940;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

Qy 1 ATGCTTCTGTGGGCATCTTAACCTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
Db 472 ATGCTTCTGTGGGCATCTTAACCTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 531
Qy 61 CCAGAGGGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGCCGCCGCTAC 120
Db 532 CCAGAGGGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGCCGCCGCTAC 591
Qy 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCCGGGACTCAATTTTTCAGATCAGCATTT 180
Db 592 TACTGGCGGGTCCCGAGGACTCCGGGGATCCGGGACTCAATTTTTCAGATCAGCATTT 651
Qy 181 CAGGAGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 240
Db 652 CAGGAGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 711
Qy 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCAGGGGCTTTTCAGACCTGGCAGCG 300
Db 712 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCAGGGGCTTTTCAGACCTGGCAGCG 771
Qy 301 TGCTTCTATTCTGGGACGTGAACCGCGGAGCGGACTCTGCTCGCTCTGAGCGCTGTCG 360
Db 772 TGCTTCTATTCTGGGCGGTGAACCGCGGAGCGGACTCTGCTCGCTCTGAGCGCTGTCG 831
Qy 361 GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC 420
Db 832 GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC 891
Qy 421 AATGCTAGCGCGCGGGGGGAGCGGACACAGCAGGGGCGACACCTTCTCAGGCGCGG 480
Db 892 AATGCTAGCGCGCGGGGGGAGCGGACACAGCAGGGGCGACACCTTCTCAGGCGCGG 951
Qy 481 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGCTGG 540
Db 952 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGCTGG 1011
Qy 541 AACCCCGCATCTTACGGGCCCTTGGACCTTAAAGCCCGCGGGGCGGCTTCCGGGAG 600
|||||

Db 1012 AACCCCGCATCTTACGGGCCCTTGACCCCTTAAAGCCCGCGGGGCTTCGGGAG 1071
Qy 601 AGTCGTAGCGCGGAGGTCTGGCGCGCAAGCGCTTTCGTCTCTATCCCGGGTACGTG 560
Db 1072 AGTCGTAGCGCGGAGGTCTGGCGCGCAAGCGCTTTCGTCTCTATCCCGGGTACGTG 1131
Qy 661 GAGACGCTGGTGTGTCGCGGACGAGTCAATGGTCAAGTTCCACGGCGCGGACCTGGAACAT 720
Db 1132 GAGACGCTGGTGTGTCGCGGACGAGTCAATGGTCAAGTTCCACGGCGCGGACCTGGAACAT 1191
Qy 721 TATCTGCTACGCTGTGTCGAACGGCGGCGACTCTACCGCCATCCAGCATCTCAAC 780
Db 1192 TATCTGCTACGCTGTGTCGAACGGCGGCGACTCTACCGCCATCCAGCATCTCAAC 1251
Qy 781 CCATCAACATCGTTGTGTCAAGGTGCTTCTTAGAGATCGTACTCCGGGCGGCAAG 840
Db 1252 CCATCAACATCGTTGTGTCAAGGTGCTTCTTAGAGATCGTACTCCGGGCGGCAAG 1311
Qy 841 GTACCGGCAATCGCGGCCCTGACGCTGGCGCAACTTCTGTGCTGGCAGAAAGCTGAAC 900
Db 1312 GTACCGGCAATCGCGGCCCTGACGCTGGCGCAACTTCTGTGCTGGCAGAAAGCTGAAC 1371
Qy 901 AAAGTGAGTGACAAGACCCCGAGTACTGGGACACTGCCATCCTCTTACCAGGCAAGAC 960
Db 1372 AAAGTGAGTGACAAGACCCCGAGTACTGGGACACTGCCATCCTCTTACCAGGCAAGAC 1431
Qy 961 CTGTGTGGAGCCACCACCTGTGACACCTGGCATGGCTGTGGTGGTACCATGTGTGAC 1020
Db 1432 CTGTGTGGAGCCACCACCTGTGACACCTGGCATGGCTGTGGTGGTACCATGTGTGAC 1491
Qy 1021 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGCTTCCATCAGCCTTCCACCATGSCC 1080
Db 1492 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGCTTCCATCAGCCTTCCACCATGSCC 1551
Qy 1081 CAGGAGCTGGGCGAGCTGTCAACATGCCCATGCCATGCAATGTGAAGTCTGTGAGGAGTG 1140
Db 1552 CAGGAGCTGGGCGAGCTGTCAACATGCCCATGCCATGCAATGTGAAGTCTGTGAGGAGTG 1611
Qy 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGGACCTCATCCAGATCGAGCTGCC 1200
Db 1612 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGGACCTCATCCAGATCGAGCTGCC 1671
Qy 1201 AACCCCTGGTACGCTGCTGCCATCATCACCGACTTCTCTGGACAGCGGCGACGGT 1260
Db 1672 AACCCCTGGTACGCTGCTGCCATCATCACCGACTTCTCTGGACAGCGGCGACGGT 1731
Qy 1261 GACTGCTCTCTGGACAAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTCGCGGGGCGC 1320
Db 1732 GACTGCTCTCTGGACAAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTCGCGGGGCGC 1791
Qy 1321 AGCTACACCTCTGAGCCAGCAGTGGCGAGTGGCTTTTGGCGTGGGCTCCAAAGCTTCT 1380
Db 1792 AGCTACACCTCTGAGCCAGCAGTGGCGAGTGGCTTTTGGCGTGGGCTCCAAAGCTTCT 1851
Qy 1381 TACATGCACTACTGCACCAAGCTGTGTGACACCGGAGGCGCAAGGACAGATGCTGTC 1440
Db 1852 TACATGCACTACTGCACCAAGCTGTGTGACACCGGAGGCGCAAGGACAGATGCTGTC 1911
Qy 1441 CAGACCCGCCACTTCCCTGGCGGATGGCACCAAGCTGTGGCGAGGCGCAAGGACAGATGCTGTC 1500
Db 1912 CAGACCCGCCACTTCCCTGGCGGATGGCACCAAGCTGTGGCGAGGCGCAAGGACAGATGCTGTC 1971
Qy 1501 AAAGGGGCTCGTGGAGAGACACAACCTCAACAGCACAG----- 1541
Db 1972 AAAGGGGCTCGTGGAGAGACACAACCTCAACAGCACAGCGCTCCGACTGACATCAT 2031
Qy 1542 -----GGTGGATGGT 1551
Db 2032 TCTCAAAACAACCTCTTATTAAAGGCTACCAATGGCCTGCACACTACACAGGTTGGATGGT 2091
Qy 1552 TCCTGGGCGAAATGGGATCCCTATGGCCCTGTCTCGGGCACATGTGTGGGGCGTGCAG 1611
Db 2092 TCCTGGGCGAAATGGGATCCCTATGGCCCTGTCTCGGCACATGTGTGGGGCGTGCAG 2151
|||||

QY	1	ATGCTTCTGCTGGGCATCCTAACCTTGCTCTTCGCCGGCGCAACGCTGGAGGCTCTGAG	60
Db	1	ATGCTTCTGCTGGGCATCCTAACCTTGCTCTTCGCCGGCGCAACGCTGGAGGCTCTGAG	60
QY	61	CCAGAGCGGAGGTAGTCGTTCCCATCCGACTGCAGCCGGACATTAACGCGCGGGCTAC	120
Db	61	CCAGAGCGGAGGTAGTCGTTCCCATCCGACTGCAGCCGGACATTAACGCGCGGGCTAC	120
QY	121	TACTGGGGGGTCCGAGGACTCCGGGGATCAGGCACTCATTTTTCAGATCACAGCATTT	180
Db	121	TACTGGGGGGTCCGAGGACTCCGGGGATCAGGCACTCATTTTTCAGATCACAGCATTT	180
QY	181	CAGGAGGACTTTTACCTACACCTGACCGCGATGCTCAGTTCTTGGCTCCCGCCTTCTCC	240
Db	181	CAGGAGGACTTTTACCTACACCTGACCGCGATGCTCAGTTCTTGGCTCCCGCCTTCTCC	240
QY	241	ACTGAGCATCTGGCGTCCCGCTCCAGGGGCTCACGGGGGCTCTTCAGACCTGCGAGCG	300
Db	241	ACTGAGCATCTGGCGTCCCGCTCCAGGGGCTCACGGGGGCTCTTCAGACCTGCGAGCG	300
QY	301	TGCTTCTATTCTGGGACGTGAACCGCAGCCGACTCGTTCGCTGCTGTGAGCCTGTGC	360
Db	301	TGCTTCTATTCTGGGACGTGAACCGCAGCCGACTCGTTCGCTGCTGTGAGCCTGTGC	360
QY	361	GGGGGGTCCGCGGAGCCTTTGGCTACCGAGGCGCCGAGTATGTCATTAGCCCGCTGCC	420
Db	361	GGGGGGTCCGCGGAGCCTTTGGCTACCGAGGCGCCGAGTATGTCATTAGCCCGCTGCC	420
QY	421	AATGCTAGCGGCGCGGGCGGCGAGCGCAACAGCGAGGGCGACACCTTCTCCAGCGCGG	480
Db	421	AATGCTAGCGGCGCGGGCGGCGAGCGCAACAGCGAGGGCGACACCTTCTCCAGCGCGG	480
QY	481	GGTGTCCGGCGGGCGCTTCGGAGACCCACCTCTCGCTGCGGGGTGGCCTCGGGGTGG	540
Db	481	GGTGTCCGGCGGGCGCTTCGGAGACCCACCTCTCGCTGCGGGGTGGCCTCGGGGTGG	540
QY	541	AACCCGCCATCTACGGGCGCTTGGACCCCTTACAAGCCGCGGGCGGGGCTTCGGGGAG	600

Query Match	22.2%;	Score 634.4;	DB 9;	Length 2184;
Best Local Similarity	60.6%;	Pred. NO. 5.5e-161;		
Matches 1097; Conservative	0;	Mismatches 701;	Indels 12;	

[illegible]

[illegible]

RESULT 8

```

RESULTS
US-10-097-580-13
; Sequence 13, Application US/10097580
; Publication No. US20030032168A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiichi
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; composition and method of immunologically analyzing human AD
; FILE REFERENCE: Q37092
; CURRENT APPLICATION NUMBER: US/10/097,580
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 05/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 2184
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(2184)

```

US-10-097-580-13

Query Match 22.2%; Score 634.4; DB 9; Length 2184;
Best Local Similarity 60.6%; Pred. No. 5.5e-161;
Matches 1097; Conservative 0; Mismatches 701; Indels 12; Gaps 3;

```
Qy 631 AAGGTTTCGTGCTATPCCGCGGTACGTGGAGACGCTGGTGGTCCGGACGAGTCAATG 690
Db 31 AAGCGATTGTGTCACGCGCCGCTTATGTGGAACACCATGCTGTAGCTGACCAAGTCCATG 90

Qy 691 GTCAAGTTCCACGGCGGACCTGGACATATCTGCTGACGCTGCTGCTGCAAGCGGGG 750
Db 91 GCGGACTTCACGGCAGCGGTCTAAAGCATTAACCTTTCTAACCCCTGTCTTGGTGGCAGC 150

Qy 751 CGACTCTACCGCCATCCACGATCCTCAACCCCATCAACATCGTTGTGGTCAAGTGTG 810
Db 151 AGTTTTACAGCATCCACGATAGGAATTCATAGCTGGTGGTGGTGAAGATCTTG 210

Qy 811 CTCTTTAGAGATCGTGACTCCGGGCCAAGGTCAACGGCAATCGGCGCCCTGACGCTGCC 870
Db 211 GTCATATACGAGGAGCAGAAAGGACCAAGAGTTACCTCCAATGACGCTCTCACCCCTCG 270

Qy 871 AACTTCTGCGTGGCAGAGAGAGCTCAACAAAGTGAGTGACAAGCACACCTCCGAGTACTG 930
Db 271 AATTTCTGACGTGGCAGAAACACACACAGCCCGAGTGACCGGATCCAGAGCACTAT 330

Qy 931 GACACTGCCATCCTTTACACGAGCAGGACCTGTGTGGAGCCACACCTGTGACACACCTG 990
Db 331 GACACTGCAATCTGTTCACACAGAGGATTTATGTGGCTCCACACAGCTGTGACACTTC 390

Qy 991 GGCATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGAGCTGCTGTCATAGGAG 1050
Db 391 GGAATGTCAGATGTGTGAACCGTATGTGACCCAGCAGGAGCTGCTCAGTCAATAGAAT 450

Qy 1051 GATGGCTTCCATCAGCTTCCACACTGCCAGAGCTGGGCCAGCTGTTTCAACATGCC 1110
Db 451 GATGTTTGAAGTGGCTTCCACAGCCCATGATGTTGGGCCATGTGTTTAACTGGCG 510

Qy 1111 CATGACAAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCCACACCATGATG 1170
Db 511 CACGATGATGTAAGCACTGTGCCAGCTTGAATGGTGTGAGTGGCGATTTTCATCTGATG 570

Qy 1171 TCCCGGACCTCATCCAGATGACCGTGGCCAAACCCCTGTGTCAGCTGCAAGTGGCCATC 1230
Db 571 GCCTCGATGCTTCCAGCTTAGACATAGACCGCCCTGGTTCACCTTGCATGGCTACATG 630

Qy 1231 ATCACCAGCTTCTGGACAGCGCAGCTGACTGCTCTGACCAAGCTCAGCAAGCC 1290
Db 631 GTCAGTCTCTTAGATATATGACAGGGGATGTTGTATGACAAAGCTCCAGATCCA 690

Qy 1291 ATCTCCCTGCCGAGGATCTGCGGGGGCCAGCTACACCTGAGCCAGAGTGGAGCTG 1350
Db 691 ATCAAGCTCCCTCTGATCTTCCGGTACCTTTGATGATGCCAACCGCCAGTGTCACTTT 750

Qy 1351 GCTTTTGGCTGGCTCCAGCCCTGCTTACATGACGTAC---TGCACCAAGCTGTGG 1407
Db 751 ACATTGCGAGAGGAATCCAGACACTGCCCTGATGACGACGACATGTACTACCTGTGG 810

Qy 1408 TGCACCGGAGGCAAGGACAGATGCTGTGCGACAGCCGCACTTCCCTTGGGCGCAT 1467
Db 811 TGCAGTGGCACTCCGCTGGCTTACTGTGTGCGCAACAAACATCTCCCTTGGCGAGAT 870

Qy 1468 GGCACAGCTGTGGCGAGGGAGCTGTGCTCCAAAGGGCGCTGGTGGAGACACAC 1527
Db 871 GGCACAGCTGTGGAGAGGGAGTGTGTGAGTGGCAAGTGCCTGGAACAGACAGAC 930

Qy 1528 CTCACAGACACAG-----GGTGGATGGTTCCTGGGCCAAATGGGATCCCTATGGCCC 1581
Db 931 ATGAAGCATTTTGTACTCTCTGTTTCAATGGAAGCTGGGGACCATGGGACCTGGGGAGAC 990

Qy 1582 TGCTCGGCGACATGTGGTGGGGCGTGCAGCTGGCCAGGAGGAGTGCACCAACCCAC 1641
Db 991 TGCTCAAGAACCTTGCTGGTGGAGTTCAATACACATGAGAGATGTGACACCCCGATC 1050
```

```
Qy 1642 CCTGCCAACGGGGGCAAGTACTGCGAGGAGTGAAGGTGAATPACCGATCCTGCAATCTG 1701
Db 1051 CCAAGAAACGGGAGGAGTACTGTGAAGCAACAGAGTCCGCTACAGTCTCTGTGAATC 1110

Qy 1702 GAGCCCTGCCCCAGCTCAGCCTCCGGGAAGAGCTTCGGGAGGAGGAGTGTGAGGCTTTC 1761
Db 1111 GAGGACTGTCCAGACATAA---CGGAAAAACGTTTACAGAGAGGAGTGGGAGTCCGAC 1167

Qy 1762 AACGGCTACAACACAGCAACCAACGCGTCACTCTCGCGGTGGCATGGGTGCCCAAGTAC 1821
Db 1168 AATGATTTTCCAAAGCTTCTTTGGGAATGAGCCCACTGTAGAGTGCACACCCCAAGTAC 1227

Qy 1822 TCCGGCGTGTCTCCCGGGGCAAGTGAAGCTATCTGCGGAGCCATGCGCACTGCTAC 1881
Db 1228 GCGGCGCTCTCGCCAAAGGACAGGTGCAAGCTCACTCTGTAAGGCCAAAGGCATTTGGTAC 1287

Qy 1882 TTCTATGTGCTGGCCACCAAGGTGTGCGACGCGTGTCTCTCTGACTGCTCACTCC 1941
Db 1288 TTTTTCGCTTTACAGCCCAAGGTGTAGTAGGCACTCCCTGTAGTCCAGACTCTACCTCT 1347

Qy 1942 GTCTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGTATGGGAACCTGGGCTCCAAGAA 2001
Db 1348 GTCTGTGTGCAAGGCAAGTGTGTAAGCTGCTGTGTATGCTCATATAGACTTCCAAAGAA 1407

Qy 2002 AGATTTCGACAGTGTGGGGTGTGTGGGGGAGACAATTAAGAGCTGCAGAAGGTGACTGGA 2061
Db 1408 AAGTTTGAATGATGTGGCGTTTGTGGAGGAACGTTTCCACATGCAATAGACTTCCAAAGAA 1467

Qy 2062 CTTCTTCAACAGCCCATCATGCGTGTACAAATTTCTGTGTGGTGGCCATCCCGCAGGCGCTCA 2121
Db 1468 ATAGTCACTAGTACAGACCTGGGTATCATGACATTTGTCACAAATCTCTGCTGGAGCCACC 1527

Qy 2122 AGCATCGACATCCCGCAGCGGTTTACAAAGGCGTGTATCGGGGATGACAACTACTCGCT 2181
Db 1528 RACATTGAAGTGAACATCGGAATCAAAAGGGGTCCAGAAACAATGGCAGCTTTCTGGCT 1587

Qy 2182 CTGAAGAACAGCCAGGCAAGTACTGCTCAACGGGCAATTTCTGTGGTGTGCGCGTGGAG 2241
Db 1588 ATTAGAGCGGTGTGTGTACCTATATTTCTGAATGGAATTTCACTCTGTCCACACTAGAG 1647

Qy 2242 CGGGACCTGGTGTGAAGGCGAGTCTGCTGCGGTACAGCGGCGGCGACAGCGGTGGAG 2301
Db 1648 CAAGACCTCACTACAAAGGTACTGCTTTAAGTACAGTGGTCTCTCGGCTGCGTGGAA 1707

Qy 2302 AGCCTGCAAGGCTTCCCGGCCCATCTGGAGCGCGTGCAGCGGTGGAGGTCTCTCGTGGGG 2361
Db 1708 AGAATCCGCGAGCTTTAGTCCACTCAAGAAAGCCCTTAACCATCCAGGTTCTTTATGTAGGC 1767

Qy 2362 AGATGACACCGCCCGGTCGCTACTCTCTCTATCTGCCCCAAAGAGCCCTCGGAGGAC 2421
Db 1768 CATGCTCTCCGACCCCAAAATTAATTCACCTACTTTTATGAAGAGAGACAGATCATTC 1827

Qy 2422 AAGTCTCTCTC 2431
Db 1828 AACGCCATTC 1837
```

RESULT 9

US-09-445-023A-13

; Sequence 13, Application US/09445023A

; Patent No. US20020119167A1

; GENERAL INFORMATION:

; APPLICANT: Hirose, Kunitaka

; APPLICANT: Inoguchi, Eiichi

; APPLICANT: Hakoizaki, Michinori

; APPLICANT: Ishioka, Keiko

; APPLICANT: Ishida, Yukako

; APPLICANT: Matsushima, Kouji

; APPLICANT: Kuno, Kouji

; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical

; TITLE OF INVENTION: composition and method of immunologically analyzing human ADA

; FILE REFERENCE: Q57092

QY 1378 CCTATACATGAGTAC---TGCACCAAGCTGTGTGTGCACCGGGAAGCCCAAGGACAGATG 1434
Db 1957 CCCGATCCAGCAGCAGATGTAGCACTTGTGTGTACCGGCACCTCTGTGTGGGTGCTG 2016
QY 1435 GTGTGCAGACCCGACATTCCTCCCTGGCGGATGAGCAGAGTGTGGCAGGCAAGCTC 1494
Db 2017 GTGTGTCAAAACAACTTCCCGTGGCGGATGAGCAGAGTGTGGAGAGGAATGG 2076
QY 1495 TGCTCAAAAGGCGCTGCGTGGAGAGACAACTCAACAAG---CACAGGTGGAT 1548
Db 2077 TGTATCAACCGAAGTGTGTGAACAAACCGACAGAAAGCATTTGTATACGCTTTTCAT 2136
QY 1549 GGTTCCTGGCCAAATGGATCCTATGGCCCTCTGCGGCACATGTGTGGGGCGCTG 1608
Db 2137 GGAAGCTGGGAATGTGGGGCGCTGGGGAGACTGTTCGAGAACGTGCGGTGGAGGATC 2196
QY 1609 CAGTGGCCAGGAGGAGTGCACCAACCCCTGCGCCACCGGGGCAAGTACTCGGAG 1668
Db 2197 CAGTACACGATGAGGAATGTGAACCCAGTCCCAAGAAATGGAGGAAGTACTGTGA 2256
QY 1669 GGAGTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCGCCAGCTCAGCCTCCGGA 1728
Db 2257 GCGAAGCAGTGGCTACACATCTGTAACTTGAAGACTGTCCAGAC---AATAATGA 2313
QY 1729 AAGAGCTTCGGGAGGAGCAGTGTGAGGCTTTCAAGGGCTTACCAACACAGCACCACCG 1788
Db 2314 AAAACCTTTAGAGAGGAACAATGTGAAGCACACAACAGTGTTCAAAAGCTTCTCTTGG 2373
QY 1789 CTCACTCTCGCGGTGGATGGTGCCCAAGTACTCGCGCTGTCTCCCGGACAAAGTGC 1848
Db 2374 AGTGGGCTCGGGTGAATGGATTCCAAGTACGCTGGCGCTCTCACCACAAAGGACAGTGC 2433
QY 1849 AAGCTCATCTGCCGAGCAATGGCAGTGGCTACTTCTATGTGTCGACCAACAGGTGGT 1908
Db 2434 AAGCTCATCTGCCAAGCAAGCATTTGGCTACTTCTCTGCTTTGACGCCAAGTGTGA 2493
QY 1909 GACGGCAGCTGTGCTCTCTGACTCCACTCCGCTGTGTGTCAAGCAAGTGCATCAAG 1968
Db 2494 GATGGTACTCCTATGATAGCTCCTCAAGAGATTCACCTCTGTGTGCAAGGACAGTGTGAAA 2553
QY 1969 GCTGGCTGTGATGGAACCTGGGCTCAAGAAGAGATTCCACAGTGTGGGTGTGTGG 2028
Db 2554 GCTGGTGTGATCGCATATAGACTCCTCAAAAGAGTTTGTATTAATGTGTGTGTGCGGG 2613
QY 2029 GGAGACAATAGAGCTCAAGAAGTGAAGTGAAGTCTTCAACCAAGCCCATGCATGCTAC 2088
Db 2614 GGAATGGATCTACTTGTAAAAAATATCAGGATCAGTTACTAGTGCAAAACCTGGATAT 2673
QY 2089 AATTTGTGTGGCCATCCCGCAGGCGCTCAAGATCGACATCCCGCAGCGGTTAC 2148
Db 2674 CATGATATCATCAAAATTCCAACTGAGCCACCAACATCGAAGTGAACAGCGGAACCCAG 2733
QY 2149 AAAGGCTGATCGGGATGACAACTACTCTGCTCTGAAGAACGCCAAGCAAGTACTCG 2208
Db 2734 AGGGATCCAGGAACATAGCGAGTTTCTTGGCCATCAAGCTGCTGTATGACACATATAT 2793
QY 2209 CTCACCGGCATTTCTGTGTGTGTCGGCGGTGAGCGGACCTGTTGTGTGAAGGCACTGTG 2268
Db 2794 CTTAATGTGTACTACTTTGTCCACCTTAGAGCAAGACATATATGACAAAGGTGTGTC 2853
QY 2269 CTGGGTACAGCGGACGGGACAGCGGTGAGAGCCTGACAGCTTCCCGGCCCATCTCTG 2328
Db 2854 TTGAGGTACAGCGGCTCTCTGCGGCATTTGGAAGAAATTCGCAGCTTTAGCCCTCTCAA 2913
QY 2329 GAGCGCTGACCGTGGAGGTCTCTCTCGTGGGAAGATGACACGCGCCCGGGTCCGCTAC 2388
Db 2914 GAGCCCTTGACCATCCAGGTCTTACTTGTGGGCAATGCCCTTCAGCCTAAATTAATATAC 2973
QY 2389 TCCTTCT 2395
Db 2974 ACCTACT 2980

RESULT 14
US-09-918-171A-8
; Sequence 8, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 3638
; TYPE: DNA
; ORGANISM: Mus musculus ADAMTS-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (278)..(2992)
; NAME/KEY: misc_feature
; LOCATION: (3636)
; OTHER INFORMATION: n = T
US-09-918-171A-8

Query Match 20.7%; Score 590.4; DB 10; Length 3638;
Best Local Similarity 57.3%; Pred. No. 4.2e-149;
Matches 1334; Conservative 0; Mismatches 906; Indels 90; Gaps 11;

QY 157 CTCATTTTTCAGATCACAGCATTTTCAGAGGACTTTTACCTACACCTGACGCCGGATGCT 216
Db 440 CTCGCTTCCACCTCTCCGCTTGGCCAGGCTTCGTGCTGGCCTGCGCCTGACGCC 499
QY 217 CAGTTCCTGGCTCCGCTTCTCCACTGAGCATCTGGCGTCCCTCCAGGGCTCACC 276
Db 500 AGCTTCCTGGCGCCGGAATTCAGATCAGGCGCTCGGGGCTCGAGCGCGCG 556
QY 277 GGGGGCTCTTCAGACCTCGAGCTGCTTCTATTCTGGGGAGCTGAACCGCGAGCGGAC 336
Db 557 GGGGGAGCGCGGACTGGTGGCTGCTTCTCTGACACAGTGAATGAGAGACGGGAG 616
QY 337 TCGTTGCTGTGTGTGAGCTGTGCGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCC 396
Db 617 TCGCTGGCGCGATGAGTGTGTGCGGGGCTGGAGCGCTGCTTCTGTCGACGCGAG 676
QY 397 GAGTATGCTATTAGCCCGCTGCCAATGCTAGCGCGCGCGGGCGCAGCAACAGCCAG 456
Db 677 GAGTTCACCATCCAGCCAGGCGCTGGGACTCCCTGGACCGCTCATCGCCTCGAG 736
QY 457 GCGGCACACCTTCTCAGCGCGCGGGGTGTTCGGCGGGGCTTCGGGAGA----CCCCAC 512
Db 737 CGCTGGGGCGGGGACAGCGCGCGGAGACCCCGGGCTCGCTGCGCGGAAGTTTCCCC 796
QY 513 CTCTCGCTGCGGGGTGGCTCTGGGCTGGAAACCCCGCATTCCTACGGGCGCTGGACCTTA 572
Db 797 CTCCCTCAAGGACTGGAGTGGGAGTGGATGGTAAATGGGACAGGACAGAGAGAAAT 856
QY 573 CAAGCGCGCGGGGCTTCGGGGAGAGTGTAGCGCGCGCGCTCTGGG----- 624
Db 857 GACAACGAAGAGAGCAAGAACAGCAGGAGGGGTGTCTCAAGAGACAGAGACTCC 916
QY 625 -----CGCGCAAGCGCTTCGTGTCTATCCG 651
Db 917 CGCAAGTCCACACCCCTTCGATCCAAAATAAGCAAGAGAGTGTGTGTCAGGCT 976
QY 652 CGGTACGTGGAGCGCTGTGTGTCGGGACGAGTCAATGTTCCAGCGCGCGGAC 711
Db 977 CGCTCGTGGAAACACTTCTGTGTGCTGATGCTCATGTGCTTCTATGGGACCGAC 1036
QY 712 CTGGAACATATCTGTGACGCTGTGGCAACGCGCGCGGCTCTTACCGCATCCAGC 771

Db	2111	AATGCCTACAACCA--CACTGACCTGGATGGGAATTTCTCTCAGTGGGTCCCCAAGTAT	2167
QY	1822	TCCGGCGTGCTCCCCGGGAGCAAGTCATCTGCCAGACCAATGGCACTTCGCTAC	1881
Db	2168	TCAGGAGTGTCCCCCGAGACCGATGACAGTGTTTTGAGAGCCGCTGGGAGGACGTGAG	2227
QY	1882	TTCTATGTGCTGGCACCAAGGTGGTGGACGGCAGCTGTGCTTCCTGACTCCACCTCC	1941
Db	2228	TTCAAGGTGTTGAAGCTAAGGTGATCGATGGCACTCTGTGTGGACCGGATACCTGTGCC	2287
QY	1942	GTCGTGTCCAAAGCAAGTCATCAGGCTGGCTGTGATGGGAACCTGGGCTCCAGAAAG	2001
Db	2288	ATCTGCGTGGGGGGCAATGTGTTAAGGCTGGCTGTGACCATGTGGTGAAGTCACTCAAG	2347
QY	2002	AGATTCACAAAGTGTGGGTGTGTGGGGGAGACAATAAGAGCTGCAAGAAGTGATCTGGA	2061
Db	2348	AAGCTGGACAAATGTGGGTGTGTGGGGCAAGGCACCTGCCTGTAGGAAGATCTCCGGT	2407
QY	2062	CTCTTCACCAAGCCCATGCAATGCTACAAATTTCTGTGTGGCCATCCCGCAGGCGGCTCA	2121
Db	2408	TCCTTTCACCCCTTCAGTTATGGCTACAAATGACATGTGCACCATCCCAAGCTGGTCCACA	2467
QY	2122	AGCATGCACATCCCGCAGCGCGTTACAAGAGGCTGATCGGGGATGACAACTACCTCGCT	2181
Db	2468	AACATTGATGTGAACAGCGGAGTACCCAGGGGTCAAGAACGCGGACGCTACTCGCGG	2527
QY	2182	CTGAAGAACGCCAAGCAAGTACCTGCTCAACGGGSCATTTCTGGTGTCCGGCTGTGGAG	2241
Db	2528	CTGAAGACAGCAATGGCAGTACCTGCTCAATGTGAACCTGGCCATCTCTGCCATAGAG	2587
QY	2242	CGGACCTGTGTGTGAAGGCCATGCTGCTCGCGGTACAGGGGACGGGGACAGCGGTGGAG	2301
Db	2588	CAAGACATCTTTGTGAAGGGGACCATCTCTGAAGTACAGTGGCTCCATGGCTTACGCTGGAG	2647
QY	2302	AGCCTGCAGGCTTCCCGGGCCATCTCTGGAGCGGCTGACCGTGGAGTGCTCTCGCTG--	2358
Db	2648	CGGCTGCAGAGCTTCCAGGCCCTGCCTGAGCCTTACAGTACAGCTCCTGACTCTGTCT	2707
QY	2358	GGGAAGATGACAGCGGCCGGGTCCGCTACTCCTTCTATCTGCCCAAGA	2408
Db	2708	GGTGAAGTCTTCCCTCAAAGTACAGATATACCTCTCTTTGTCCCAATGA	2757

DEBIT 15

US-10-174-390-331

: GENERAL INFORMATION.

: APPLICANT: Baker, Kevin P

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers

APPLICANT: Goddard, A

; APPLICANT: Godowski,

; APPLICANT: Gurney, Au

; APPLICANT: Pan, James

; APPLICANT: Smith, Vic

; APPLICANT: Watanabe,

; APPLICANT: Wood, Will

; APPLICANT: Zhang, Zeng

; TITLE OF INVENTION: S

; TITLE OF INVENTION:

; FILE REFERENCE: P3430

; CURRENT APPLICATION N

; CURRENT FILING DATE:

; Prior application removed

; NUMBER OF SEQ ID NOS:
: CEO ID NO 351; SEQ ID NO 351
; LENGTH: 4407

LENGTH: 440/
TYPE: DNA

TYPE: DNA
ORGANISM: Homo sapiens

ORGANISM: HOMO SAPIENS
NC-10-174-590-351

DEC 4/7 07 30

Query Match

1103 MW 7 7000X

Best Local Similarity 60.3%; Pred. No. 1e-146;
Matches 1092; Conservative 0; Mismatches 694; Indels 24; Gaps 7;

QY 609 CCGCGCAGGTCTGGCGCGCCCAAGGCTTGTGTTATCTCCCGGTAGCTGGAGACGCT 668
DB 1027 CAGCCCCAGACCCGAGAGCCCAAGGCTTGTCTTCACTGAGTAGATTCTGGAGACAT 1086
QY 669 GGTGGTGGCGGAGCTCAATGTCAAGTTCACGCGCGGACCTGGGAACATATCTGCT 728
DB 1087 GGTGGTGGCAGATGACAAGATGGCCCATCTCACGGTGGGGCTAAAGCGCTACTGCT 1146
QY 729 GAGCTGCTGGCAACCGCGCGGAGCTCTACCGCCATCCAGCATCTCTCAACCCCATCAA 788
DB 1147 AACAGTGATGGCAGCAGCAGCCAAAGGCTTCAAGCAACCCAAAGCATCCGCAATCTGTCA 1206
QY 789 CATCGTTGGTCAAGGTGCTGCTTCTTAGAGATCTGACTCCGGCCCCAAGGTCAACGG 848
DB 1207 CTTGGTGGTGACTCGGCTAGTGATCTGGGGTTCAGCGAGGAGGGCCCTAAGTGGGGC 1266
QY 849 CAATGGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAAGCTGAACAAAGTGAG 908
DB 1267 CAGTGTGCCAGACCCCTGGCAGCTTCTGTGCTGGCAGCGGGGCTCAACACCCCTCA 1326
QY 909 TGACAGCACCCGAGTACTGGACACTGCCATPCTCTTCAACAGGACGAGCTGTGTGG 968
DB 1327 GGACTGGGGCCCTGACCACTTTGACACAGCCATTCTGTATTACCCGTCAGAGCTGTGTG 1386
QY 969 AGCCACCACCTGTGACACCTGGGCTGATGTGGTACCATGTGTGACCCCAAGAG 1028
DB 1387 AGTCTCCATTTGGCAGCAGCTGGGTATGGTGTGGCAGCGCTGTGACCCGCTCG 1446
QY 1029 AAGTGTCTGTCTATTGAGGACGATGGGTTCCATCAGCTTCAACACTGCCACAGCT 1088
DB 1447 GAGCTGTGCCATTGTGGAGATGATGGGCTCCAGTCCAGCTTCACTGCTCATGAAT 1506
QY 1089 GGGCCAGCTTCAACATGCCCCCATGACATGTGAAGTCTGTGAGGAGGTGTTGGGAA 1148
DB 1507 GGGTCTATGTCTTCAACATGTCTCATGACAACTCCAAAGCCATGATCATGTTGAATGGGC 1566
QY 1149 GCTCCGAGGCAAA---CCACATGATGTCCCGACCCCTCATCCAGATCGACCGTGCCTCAAC 1205
DB 1567 TTTGAGCACCTCTCGCATGTCTATGGCCCTGTGATGGCTCATGTGGATCTTGAGAGCC 1626
QY 1206 CTGCTAGCTGAGTGTGCCATPACATCCGACTTCTGTGA---AGCGGCGAGGTGACTG 1265
DB 1627 CTGCTCCCTCGAGTCCCGCTCTATCACTGACTTCTCTGGA---AATGGCTATGGGCACTG 1686
QY 1266 CCTCTGGACCAACCCAGACGCCATCTCCCTGCCGAGGATCTCGCGGGGCCAGCTA 1325
DB 1687 TCTTTAGAACAAACAGAGGCTCCATGTGATCTGCTGTGACTTTCCCTGGCAAGGACTA 1746
QY 1326 CACCCCTGAGCCAGAGTGCAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCTTTACAT 1385
DB 1747 TGATGCTGACCGCCAGTGCAGCTGACTTTCGGGCGGACTCACGCCATTGTCCAGCT 1806
QY 1386 ---GCAGTACTGACCAAGCTGTGGTGCACCGGGAAGCCAGGACAGATGCTGTGCTCA 1442
DB 1807 GCCCGCCCTGTGCTGCCCTCTGGTCTCTGGCCACCTCAATGGCCATGATGCTGCTGCA 1866
QY 1443 GACCCGCACTTCCCTGGGCGGATGACACAGCTGTGGGAGGCAAGCTGTGCTCA 1502
DB 1867 GACCAACACTGCGCTTGGCCGATGGCACACCCCTGGGCGCCGACAGGCCCTGCAATGG 1926
QY 1503 AGGGCCCTGGTGGAGAGACACACCTTCAACAGCACAGGCT-----GGAGGTTCTCTG 1556
DB 1927 TGGTGTGCTGCTCCACATGGACAGCTCCAGGACTTCAATATCCACAGGCTGGTGGCTG 1986
QY 1557 GGCCAAATGGATCCCTATGGCCCTGCTCGGCACATGTGTGGGGGCTGAGCTGGC 1616
DB 1987 GGGTCTTGGGGACCATGGGGTACTGCTCTCGGACCTGTGGGGGTGGTGTCCAGTCTC 2046
QY 1617 CAGGAGGCTGACCAACCCACCCCTGCTGCAACGGGGCAAGTACTTGCAGGAGGTGAG 1676

DB 2047 CTCCCGAGACTGCACGAGGCTGTCTCCCGGGAATGTGGCAAGTACTGTGAGGCGCCCGC 2106
QY 1677 GGTGAATACCGATCTCTGCAATCTGGAGCCTGCCAGCTCAGCCTCCCGGAATATCTT 1736
DB 2107 TACCCGCTTCGGCTCTGCAACTGAGGACTGCCAA---CTGGCTCAGCCCTGACCTT 2163
QY 1737 CCGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAACCAACCGGCTCACTCT 1796
DB 2164 CCGCAGGAGCAGTGTGCTGCTTAAACACCGCACCGAGCTCTTCAAGAGCTTCCAGG 2223
QY 1797 CGCGTGGCATGGTGGCCCAAGTACTCCGGGCTGTCTCCCGGGGACAAGTGCAGCTCAT 1856
DB 2224 GCCCATGGACTGGTGTCTCTGCTACAGCGGTGGCCCCCAGGACGAGTGCATACTAC 2283
QY 1857 CTGCGGAGCCAAATGGCACTGGCTACTTCTATGTCTGGCACCCCAAGTGTGGACGCAC 1916
DB 2284 CTGCGAGCCCGGCACTGGCTACTACTATGTCTGGAGCCAGGGTGTAGATGGAC 2343
QY 1917 GCTGTGCTTCTCTGACTCCAGCTCCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTG 1976
DB 2344 CCCCTGTTCGCCGACAGCTCCTCGGTGTGTCTCAGGCGCATGCTCATGCTGGCTG 2403
QY 1977 TGATGGGAACCTGGCTCCCAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAA 2036
DB 2404 TGATGCACTATTGGCTCCCAAGAAAGTTTGACAAGTGTGATGTGGGAGGGACGG 2463
QY 2037 TAAGAGCTGCAAGAGTGTGACTGGACTCTTCAACCAAGCCCATGCTGCTGCTACAAT 2096
DB 2464 TCTGTGTGACAGCAGCAGTCAAGCTCCTTCAGGAATTCAGGTACGATACAACATGT 2523
QY 2097 GGTGGCCATCCCGCAGCGGCTCAAGCATCGCATCCGCCAGCGCGGTTACAAGGGCT 2156
DB 2524 GGTCACTATCCCGCGGGGCGCACCATTTCTTGTCCGGCAGCAGGGAACCCCTGGCCA 2583
QY 2157 GATCGGGATGACAACTACCTGCTCTGAAGAACAGCAAGGCAAGTACCTGCTCAACGG 2216
DB 2584 CCGGAG-----CATCTACTTGGCCCTGAACTGCCAGATGGCTCTATGCCCTCAATGG 2637
QY 2217 GCATTTCTGTGTGTGGCGGTGGAGCGGACCTGCTGGTG--AAGGCGAGTCTGC-TGCG 2273
DB 2638 TGAATACACGCTGATGCCCTCCCGCACAGATGTGTACTGCTGGGCGAGTCAAGTTCG 2697
QY 2274 GTACAGCGGCGGCGCACAGCGGTGGAGAGCTTCAGGCTTCCCGGCCCATCTCGAGCC 2333
DB 2698 CTACAGCGGGGCGCACTGACGCTTCAGAGACACTGTCAAGCCATGGGCCACTGGCCAGCC 2757
QY 2334 GTTACCGGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGGTCCGCTACTCTT 2393
DB 2758 TTTGACACTGCAAGTCTCTAGTGGCTGGCAACCCCGCAGGACACACGCTTCCGATACAGCTT 2817
QY 2394 CTATCTGCC 2403
DB 2818 CTTCTGTGCC 2827

Search completed: May 16, 2003, 00:52:32
Job time : 382 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 19:42:09 ; Search time 107 Seconds
(without alignments)
8177.092 Million cell updates/sec

Title: US-09-965-631-3
Perfect score: 2853
Sequence: 1 atgtcttctgtggcgcctct.....gcgtctctgagcgctgctga 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	626.6	22.0	4676	4	US-09-130-491-1	Sequence 1, Appli
2	617.2	21.6	4858	4	US-09-392-184-1	Sequence 1, Appli
3	611.4	21.4	3706	4	US-09-484-970B-58	Sequence 58, Appl
4	598.2	21.0	3126	4	US-09-392-184-7	Sequence 7, Appli
5	590.4	20.7	3638	4	US-09-369-364A-8	Sequence 8, Appli
6	583.2	20.4	4192	4	US-09-122-126B-1	Sequence 1, Appli
7	549	19.2	3250	4	US-09-122-126B-14	Sequence 14, Appli
8	547.4	19.2	3002	4	US-09-369-364A-1	Sequence 1, Appli
9	469.8	16.5	2625	4	US-09-369-364A-14	Sequence 14, Appli
10	409.6	14.4	5804	4	US-09-369-364A-12	Sequence 12, Appli
11	318	11.1	2114	4	US-09-130-491-7	Sequence 7, Appli
12	299.4	9.1	7739	4	US-09-369-364A-10	Sequence 10, Appli
13	245.6	8.6	5357	4	US-09-392-184-5	Sequence 5, Appli
14	238.6	8.4	3218	4	US-09-369-364A-6	Sequence 6, Appli
15	217.4	7.6	1520	4	US-09-369-364A-3	Sequence 3, Appli
16	212.8	7.5	3885	4	US-09-369-364A-16	Sequence 16, Appli
17	185.4	6.5	2450	4	US-09-491-522-2	Sequence 2, Appli
18	185.4	6.5	6692	4	US-09-491-522-1	Sequence 1, Appli
19	177.8	6.2	2450	4	US-09-491-522-9	Sequence 9, Appli
20	177.8	6.2	4580	4	US-09-491-522-8	Sequence 8, Appli
21	162.2	5.7	703	4	US-09-392-184-6	Sequence 6, Appli
22	116	4.1	3675	4	US-09-392-872-3	Sequence 3, Appli
23	116	4.1	4042	4	US-09-930-872-5	Sequence 5, Appli
24	91.6	3.2	2848	4	US-09-369-364A-4	Sequence 4, Appli
25	79	2.8	2023	4	US-09-491-522-6	Sequence 6, Appli
26	77	2.7	385	4	US-09-392-184-23	Sequence 23, Appli
27	67.6	2.4	349	4	US-09-392-184-24	Sequence 24, Appli

28	61.4	2.2	1642	4	US-09-369-364A-18	Sequence 18, Appli
29	57	2.0	1476	4	US-09-930-872-1	Sequence 1, Appli
30	55.8	2.0	1803	4	US-09-369-364A-20	Sequence 20, Appli
31	50.4	1.8	657	3	US-08-985-526-2	Sequence 2, Appli
32	50.4	1.8	1326	3	US-08-985-526-4	Sequence 4, Appli
33	49.4	1.7	716	4	US-09-276-531-117	Sequence 117, App
34	47.4	1.7	3014	2	US-08-808-982-1	Sequence 1, Appli
35	47.4	1.7	3014	4	US-09-306-902A-1	Sequence 1, Appli
36	47.2	1.7	1173	3	US-08-706-216-5	Sequence 5, Appli
37	46.4	1.6	1101	4	US-09-152-060-23	Sequence 23, Appli
38	44.6	1.6	502	4	US-09-392-184-12	Sequence 12, Appli
39	44	1.5	7452	3	US-08-592-500-1	Sequence 1, Appli
40	44	1.5	7452	3	US-08-195-006-1	Sequence 1, Appli
41	44	1.5	7452	5	PCT-US94-07644A-1	Sequence 1, Appli
42	43.6	1.5	2063	1	US-08-715-554-1	Sequence 1, Appli
43	43.6	1.5	2063	1	US-08-583-118-1	Sequence 1, Appli
44	43.4	1.5	7218	2	US-08-232-463-14	Sequence 14, Appli
45	43	1.5	20235	1	US-07-642-734C-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-130-491-1
; Sequence 1, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (460)...(3360)
US-09-130-491-1

Query Match 22.0%; Score 626.6; DB 4; Length 4676;
Best Local Similarity 57.0%; Pred. No. 1.1e-131;
Matches 1304; Conservative 0; Mismatches 914; Indels 69; Gaps 6;
QY 175 GCATTTCAGAGACATTTTACCTACACCTGACCGCGGAGCTCAGTCTTGGCTCCGCGC 234
Db 697 GCCTTTACACGACGCTGGATCTGGAGCTGCGCGCCGACAGCAGCTTTTGGCGCGCGC 756

QY 235 TTCTCCACTGACGATCTGGCGCTCCCTCCAGGGCTCACCGGGGGCTC---TTCAGAC 291
Db 757 TTCACGCTCCAGAACGCTGGGGCGCAATCGGGTCCGAGACGCCGCTTCCGGAAACAGAC 816
QY 292 CTGCGACGCTGCTTCTATTCTGGGACGCTGAGCGCGGACGCTGCTTGGCTGTGTG 351
Db 817 CTGGCGCACTGCTTCTACTCCGSCACCGTGAAATGGGATCCAGCTCGGCTCGCGGCTC 876
QY 352 AGCTGTGGGGGGGCTCCGCGAGCCTTTGGCTACCGAGGCGCGAGTATGTCATTAGC 411
Db 877 AGCCTTGGAGGGCGCTGCGCGCGCTTCTACTCTGCTGGGGAGCGGCTTATTCATCCAG 936
QY 412 CGCTCCCAATGCTAGCGCGC-CGGCGGCGGACGCGCAACACCCAGGGGGCGACACTTCT 470
Db 937 CGCTGCGCGCGCGACGAGCGCTCGCCACCGCGCGCGCGGAGAGAGCGCGCGCA 996

QY 471 CCAGCCGCGGGTGTTCCTGGGGGGGCTTCCGAGACCCCACTCTCTGCTGGGGGTGC 530
Db 997 CCACACAGTTCCACCTCTCTGGGGGAATCGCAGGGGCGAGTAGGCGGCACGTCGCG 1056
QY 531 CTCGGGCTGGAACCCCGCATCTAGGGGCTTGGACCTTACAGCCCGCGGGGGG 590
Db 1057 GTCGTGAGACAGAGCCCGGGCGACTGGGAAGCGGAGAGAGAGAGAGAGAGAG 1116
QY 591 CTTTCGGGAGAGTCGTAGCCGGCGCAGTCTGGGCGC----- 627
Db 1117 ACTGAGGGCGAGACGAAGGGCTCAGTGTGCGCCGAGAGACCCGGCACTGCAAGCGGTA 1176
QY 628 -----GCCAAGCGTTTCGTGTCTATCCTCCGGGTAC 657
Db 1177 GGACAGCCACAGAACTGGAACATAGAAGAAAGCGAATTTGTGCCAGTCAACCGCTAT 1236
QY 658 GTGAGAGCGTGTGTGGGCGGAGAGTCAATGGTCAAGTTCCAGGGCGCGACCTGGNA 717
Db 1237 GTGGAACCATGCTTGTGGCAGACCAAGTGTGGCAGAAATTCACGGCAGTGGTCTAAG 1296
QY 718 CATTATCTGCTGACGCTGTGGCAACGGCGGCGACTCTACGCCATCCACAGCATCCPC 777
Db 1297 CATTACCTCTACGTTGTTTCGGTGGCAGCCAGATTTGTAACACCCACAGCATCTGT 1356
QY 778 AACCCATCAACATCTGTGTGTCAGAGTCTCTCTTTAGAGATCGTCACTCCGGGCC 837
Db 1357 AATTCACTTAGCCTGTGGTGTGGAATCTTGGTTCATCCAGATGAACAGAGGGCGC 1416
QY 838 AAGTCAACGGCAATGGGCGCTGAGCTGGCGCACTTCTGTCCCTGGCGAGAGAGCTG 897
Db 1417 GAATGACCTCCAAATGCTGCCCTCACTCTGCGGAATTTTGCAACTGGCAGAGAGCAG 1476
QY 898 AAAAAAGTGAAGCAAGCAACCCGAGTACTGGGCACTGCCATCTCTTCCACCGGAC 957
Db 1477 AACCCACCACTGACCGGATGAGAGCACTATGACACAGCAATCTTTTCACAGACAG 1536
QY 958 GACTGTGTGGAGCCACCACTGTGACACCCCTGGGATGGCTGATGGGTACCATGTGT 1017
Db 1537 GACTTGTGTGGTCCCAAGACATGTACTCTTGGGATGGCTGATGTGGAACTGTGTGT 1596
QY 1018 GACCCAGAGAGCTGCTCTGTCATTGAGGAGATGGGCTTCATCAGCTTCACCACT 1077
Db 1597 GATCCGAGAGAGCTGCTCCGTCATAGAGATGATGGTTTACAGCTCCCTTCACACA 1656
QY 1078 GCCACAGAGTGGGCGACGCTTCAACATGCCCATCAACATGTAAGTCTGTAGGAG 1137
Db 1657 GCCCATGAATTAGGCCACGTTTAAATGCCACATGATGATCAAAAGCAGTGTGCCAGC 1716
QY 1138 GTGTTGGGAGCTCGAGCCACACAGATGTCCCGACCTCATCCAGATCGACCTG 1197
Db 1717 CTTAATGGTGTGAACAGGATTCACATGATGGCGTCAATGCTTTCCAACTGGACCA 1776
QY 1198 GCCAACCCCTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1257
Db 1777 AGCAGCCTTGTGCTCTTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1836
QY 1258 GGTGACTGCTCTTGGACCAACCAAGCAAGCCATCTCCCTGCGGAGGATGTGCGGGC 1317
Db 1837 GGGGAATGTTTGTGAGCAAGGCTCAGATCCCATACAGTCCCAAGCTCCCTCCCTGGC 1896
QY 1318 GCCAGCTACACCTGAGCGAGAGTGGAGTGGCTTTGGCGTGGCTCCAAAGCCCTGT 1377
Db 1897 ACCTCGTACAGTGCACCGGAGTGGCAGTTTACATTTGGGAGGAGCTCCAAACACTG 1956
QY 1378 CTTTACATGAGTAC---TGACCAAGCTGTGTGTGACCGGAGGCGCAAGGACAGATG 1434
Db 1957 CCCGATGACCCAGCACATGTAGACCTTGTGGTGTACCGGACCTCTGTGGGTGGTGTG 2016
QY 1435 GTGTGCGAGACCCGCACTTCCCTGGGCGGATGGCACACAGCTGTGGCGAGGCGCAAGCT 1494
Db 2017 GTGTGTCAAAACCAACACTTCCCTGGGCGGATGGCACACAGCTGTGGAGAGGAAATG 2076
QY 1495 TGCCTCAAGGGGCGCTCGTGGAGAGACACACCTCAACAAG-----CACAGGTGGAT 1548

Db 2077 TGTATCAACGGCAAGTGTGTGAACAAAACCCACAGAAGCATTTTGATACGCTTTTCAT 2136
QY 1549 GGTTCCTGGGCAAAATGGGATCCCTATGTCCTGCTCGCCACATGTGTGGGGCGTG 1608
Db 2137 GGAAGCTGGGAATGTGGGGCTTGGGAGACTGTTCGAGAAGCTGCGGTGGAGAGATC 2196
QY 1609 CAGTGGCCAGGAGGAGTGCACCAACCCCTCCCAACCGGGGCAAGTACTCGGAG 1668
Db 2197 CAGTACAGATGAGGAATGTGACAAACCCAGTCCCAAGAATGGAGGAAGTACTGTGA 2256
QY 1669 GGAGTGGGTGAATACCGATCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCGGA 1728
Db 2257 GGCAAACAGTGGCTACAGATCTGTAACTTGGAGACTGTCCAGAC---AATAATGGA 2313
QY 1729 AAGAGCTTCCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTAAACACAGCAGCAACCGG 1788
Db 2314 AAAACCTTTAGAGAGAACATGTGAAGCAGACACACAGATTTTCAAAAGCTTCTTTGGG 2373
QY 1789 CTCATCTCTGCGGTGGCATGGGTCGCCAAGTACTCCCGGCTGTCTCCCGGGCAAGTGC 1848
Db 2374 AGTGGGCTGCGGTGGAATGATTTCCCAAGTACGCTGGCGTCTCACCAGAGCAGGTGC 2433
QY 1849 AAGCTCATCTCCGAGCAATGGCAGTGGCTACTTCTATGTGTGGCACCACCAAGTGTG 1908
Db 2434 AAGCTCATCTCCGAGCAATGGCAGTGGCTACTTCTGCTTTTGGAGCCCAAGTGTGA 2493
QY 1909 CAGGCGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1968
Db 2494 GATGGTACTCCATGTAGCCAGATTCACCTCTGCTGTGTGCAAGCAGAGTGTGAAA 2553
QY 1969 GCTGCTGTGATGGAGCCTGGGCTCCAAAGAGATTCGCAAGTGTGGGGTGTGGG 2028
Db 2554 GCTGCTGTGATGCGATCATAGACTCCAAAAGAGTTTGTATAATGTGTGTGTGGGG 2613
QY 2029 GGAGACATAGAGTGCAGAGGTGACTGACTCTTCAACAGCCCATGCTATGCTGCTAC 2088
Db 2614 GGAATGATGATCTGTGTAAGAAATATCAGGATCAGTTACTGTGTAAGAACTGTGATAT 2673
QY 2089 AATTTCTGTGGGCTATCCCGCAGGGGCTTCAAGCATCGACATCCCGCAGCGGGTTAC 2148
Db 2674 CATGATATCATCAATTTCCAACTGGAGCCACCAACATCGAAGTGAACAGCGGAAC 2733
QY 2149 AAAGGCTGTGCGGATGACATCTGCTGCTGAGAACACAGCAAGCAAGTACCTG 2208
Db 2734 AGGGATCCAGGAACATGGCAGCTTTCTGCCATCAAGCTGTGATGGCAGCATATAT 2793
QY 2209 CTCAGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2268
Db 2794 CTTAATGTGACTACACTTTGTCACCTTAGAGCAAGCATTTATGTAAGAGTGTGTC 2853
QY 2269 CTGCGGTACAGCGGCGCAGCGCTGAGAGCCTGAGAGCTTCCCGGCCCATCTG 2328
Db 2854 TTGAGTACAGCGCTCTCTGCGGCAATTTGAAAGAAATTCGACGTTTAGCCCTCTCAA 2913
QY 2329 GAGCGCTGAGCGTGGAGTCTCTCCGTGGGAGATGACACCCCGGCTCCGCTAC 2388
Db 2914 GAGCGCTGAGCATCCAGGTTCTTACTGGGCAATGCCCTTCGACCTTAAATTAATAC 2973
QY 2389 TCTTCT 2395
Db 2974 ACCTACT 2980

RESULT 2

US-09-392-184-1
; Sequence 1, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55

```

; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4858)
; OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
; NAME/KEY: misc_feature
; LOCATION: (1)...(4858)
; OTHER INFORMATION: n = A,T,C or G
US-09-392-184-1

Query Match      21.6%; Score 617.2; D3 4; Length 4858;
Best Local Similarity 57.0%; Pred. No. 1.5e-129;
Matches 1305; Conservative 0; Mismatches 913; Indels 70; Gaps 7;

QY 175 GCATTTCAGGAGGACTTTTACTACACCTGACGCGCGGATGCTCAGTTTGGCTCCCGCC 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 GCCTTTGACACGAGCTGGATCTGGAGCTGGCGCCGACACAGCAGCTTTTGGCGCCCGC 750

QY 235 TTCTCCACTGAGCATCTGGCGTCCCTCCAGGGGCTACCGGGGGCTC---TTGAGAC 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 TTCAGCTCCAGACGCTGGGCGCAATCCCGGTCGCGAGCGCGCTTCGGAACCGAC 810

QY 292 CTGCGACGCTGCTTATTCTGGGACGCTGAACCGCGAGCGCGGACTCTGTTGCTGCTGTG 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 CTGGCGCACTGCTTCTACTCCGGCACCCTGAATGGCGATCCACGCTGGTGGCGCCCTC 870

QY 352 AGCTGTGCGGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGGATGATGTCATTAGC 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 871 AGCTCTCGAGGGGCTGCGGGGCGCTTCTACCTGCTGGGGAGGCGTATTTTCATCAG 930

QY 412 CGCTGCGCCATGCTAGGCGC-CGGGGCGCAGCGCAACAGCAGCGCGCACACCTTCT 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 931 CGCTGCGCGCGCGCAGCGGCTGCGCACCGCGCGCTCAGGGGAGAGCGCGCGCA 990

QY 471 CCAGCGCGGGGTGTTCGGGCGGCGCTTCGCGAGACCCACCTCTCGCTCGGGGTGGC 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 991 CCATACAGTTCCACCTCTCTCGCGGGAATCGCGAGGCGACGTAGCGGCGACGTGCGGG 1050

QY 531 CTCGGGTGGAACCCCGCATCTACGGGCGCTTGACCCCTTACAGCGCGGGGCGGG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1051 GTCGTGGACGACGAGCGCGCGCTGCGCAACCGCGCTGGGAAAGCGGAGACCGAGG 1110

QY 591 CTTGGGGGAGAGTCTGAGCGGCGCAGTCTGGCGC----- 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1111 ACTGAGCGGAGGAGGCGCTCAGTGTGCGCGCGAGGACCGGCGCACTGCAAGGCGTA 1170

QY 628 -----GCCAAGCGTTTGTGTCTATCCCGCGTAC 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1171 GGACAGCCACAGGAAGCTGGAAGCATAGAAGAGCGATTTGTGTCAGTACCGGCTAT 1230

QY 658 GTGAGAGCGTGTGTGCGGAGAGTCAATGTCAAGTTCCACGCGCGGAGACTGGAA 1717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1231 GTGGAACCAATGCTGTGGCAGACAGTGTGATGCGCAATTCACGCGCGTGTCTAAAG 1290

QY 718 CATTATCTGCT-GAGCGTGTGGCAACGGCGCGGCGACTTACCGGCACTCCAGCATCT 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1291 CATTACCTTTCACAGTGTGTTTCGGTGGCAGCGAGATTTACAAACAOCCAGCATTCG 1350

QY 777 CAACCCCATCAACATCTGTGTGTCAGAGTGTCTTCTTTAGAGATGCTGACTCCGGGC 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1351 TAATTAGTGTAGCTGTGTGTGTGAGATCTTGGTCATCCAGATGAACAGAGGGGCC 1410

QY 837 CAAGGTCAACGGCAATCGGCGCTGACGCTGCGCACTTCTGTGCTGTGGAGAGAGCT 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 GGAAGTGACCTCCAACTGCTGCGGCACTTTTGCAACTTTTGCAACTGGAGAGCA 1470
```

```

QY 897 GAACAAAGTGAAGTGAAGCAAGCAAGCAAGTACTGGGACACTGCCATCTCTTCAAGCA 956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1471 CAACCCACCCAGTGACCGGATGACAGGCACTATGACACAGCAATCTTTTCAACAGACA 1530

QY 957 GGACTGTGTGGGACCAACCACTGTGACACCCCTGGGATGGCTGATGTGGGTACATGTG 1016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1531 GGACTGTGTGGGTCCAGACATGTGATCTCTTGGGATGGCTGATGTGTGGAACACTGTG 1590

QY 1017 TGACCCCAAGAGAGCTGCTCTCTGATGAGGAGATGGCTTCCATCAGCTTCAACCAC 1076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1591 TGATCCGAGCAGAAGCTGCTCCGTCATAGAGATGATGTTTACAGCTGCTTCAACCAC 1650

QY 1077 TGCCACGAGCTGGGCGACGCTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAG 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1651 AGCCATGAATTAGGCGACGCTTTAAATGCCACATGATGATGCAAGCAGTGTGCCAG 1710

QY 1137 GGTGTTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGACCTCATCCAGATCCAGC 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1711 CTTAATGTTGTTGAACAGGATTTCCACATGATGGGCTCAATGCTTTCCAACTTGACCA 1770

QY 1197 TGCCAAACCCCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1771 CAGCCAGCTTGTCTCTTCCAGTGCCTACATGATGATGATGATGATGATGATGATGATG 1830

QY 1257 CGGTGACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1831 TGGGGAATGTTGATGGCAAGCTCAGATCCATACAGCTCCAGGCGATCTCCCTGG 1890

QY 1317 CGCCAGCTACACCTGAGCAGCAGTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1891 CACCTGCTAGATGCCAACCGGCGTGCAGTTTACATTTGGGGAGGACTTCCAAACACTG 1950

QY 1377 TCCTTATACATCAGTAC---TGCCACCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1951 CCCTGATGACGCGCAGCAGATGTAGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2010

QY 1434 GGTGTGCGAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2011 GGTGTGTCAAAACAACTTCCGCTGGCGGATGCGCAGCTGTGGAGAAGGAAATG 2070

QY 1494 CTGCTTCAAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2071 GTGTATCAACGGCAAGTGTGTGAACAAACCGACAGAAAGCATTTTGTATACGCTTTTCA 2130

QY 1548 TGTCTCTGCGGCAAAATGGGATCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2131 TGAAGCTGGGGAATGCTGGGGGCTTGGGGAGACTGTTGAGAGACGCTGCTGCTGCTGCTG 2190

QY 1608 GCAGCTGGCGCAGGAGCAGTGCACCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2191 CCAGTACAGGATGAGGGAATGTGACAAACCGCTGCCAAAGATGGAGGGAAGTACTGTGA 2250

QY 1668 GGCAGTGGGTGAATACCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2251 AGCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2307

QY 1728 AAAGAGCTTCCGCGGAGGAGCAGTGTGAGGCTTTTCAACGCTTACACAGCAGCAGCAACCG 1787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2308 AAAACCTTTAGAGAGGAACATGTGAACACACACAGAGTTTCAAAAGCTTCTCTTGG 2367

QY 1788 GCTCACTGCGCGTGGCATGGGTCGCCAAGTACTCCCGGCTGTCTCCCGGCAAGTGA 1847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2368 GAGTGGGCTGCGGTGGAATGGATTTCCCAAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2427

QY 1848 CAAGCTCATCTCCCGAGCAATGGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2428 CAACTCATCTGCCAAGCAAGGCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2487

QY 1908 GGAAGGCAAGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2488 AGATGCTACTCCATGTAGCCCAAGATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2547

QY 1968 GGCTGCTGCTGAGGGAACCTGGGCTTCCAGAAAGATTCGACAAAGTGTGGGCTGCTGCTG 2027
```

```
Db 2548 AGCTGGTTGTGATCGCATCATAGACTCCAAAGAAAGTTTGATAAATGTGGTGTGGCG 2607
QY 2028 GGGAGACAATAAGAGCTGCAAGAGTGACTGGACTCTTCCACCAAGCCCATCATGGCTA 2087
Db 2608 GGGAAATGGATCTACTTGTAAAGAAATATACAGATCAGTTACTAGTGCAAAACCTGGATA 2667
QY 2088 CAATTTGTGTGGCATCCCGCAGCGCTCAAGCATCGACATCCGCTAGCGGGTTA 2147
Db 2668 TCATGATATCATCAATATCCAACTGGAGCCACCAATCGAAGTGAAGACGGGNACCA 2727
QY 2148 CAAGGCGCTGATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGCAAGTACCT 2207
Db 2728 GAGGGATCAGGACAAATGGCAGCTTCTTGCCATCAAGCTGCTGATGGACATATAT 2787
QY 2208 GCTCAAGCGCATTTGTGTGTGCGGGTGGAGCGGGACCTGGTGGTGAAGGGCAGTCT 2267
Db 2788 TCTTAATGGTGACTACACTTTTGTCCACTTAGAGCAAGACATTTATACAAAGGTGTTG 2847
QY 2268 GCTGCGGTACAGCGGCACGGGCACAGCGGTGGAGAGCCTGCAGGCTTCCTCGGCCATCCT 2327
Db 2848 CTTGAGGTACAGCGCTCCTCTCGGCATTTGGAAGAAATTCGACGCTTTAGCCCTCTCAA 2907
QY 2328 GGAGCCGCTGACCGTGAGGTCTCTCCGHHGGGAAGATGACACCGCCCGGGTCCGCTA 2387
Db 2908 AGAGCCCTTGACCATCCAGGTCTTCTACTGTGGCAATGCCCTTCGACCTAAATTAATA 2967
QY 2388 CTCCTTCT 2395
Db 2968 CACCTACT 2975
```

RESULT 3

US-09-484-970B-58
; Sequence 58, Application US/09484970B
; Patent No. 6426186

GENERAL INFORMATION:

; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484, 970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 58
; LENGTH: 3706
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 007074.1

; NAME/KEY: unsure
; LOCATION: 3634, 3638-3639, 3642-3643, 3647-3648, 3652, 3654-3658, 3664, 3674, 3681
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-58

Query Match 21.4%; Score 611.4; DB 4; Length 3706;

Best Local Similarity 60.6%; Pred. No. 2.7e-128;

Matches 1076; Conservative 0; Mismatches 686; Indels 13; Gaps 4;

```
QY 631 AAGCGTTTCGTGCTATCCCGCGGTAGCTGGAGACGCTGGTGGTTCGGGACGAGTCAATG 690
Db 183 AAGCGATTGTGTCAGTCAACCGTATGTGGAAACCATGCTTGTGGCAGACCATCGATG 242
QY 691 GTCAGTTCACGGCGGGACCTGGACATTAATCTGTGAGCTGTGCTGGCAACGGCGCG 750
Db 243 GCAGAAATCCACGGCAGTGGTCTAAAGCAATTAACCTTCTACGTTGTTTCGGTGGCAGCC 302
QY 751 CGACTCTACCGCCCATCCAGCATCCTCAACCCCATCAACATCGTTGTGGTCAAGGTGCTG 810
Db 303 AGATTGACAAACACCCACGACATTCGTAATTCAGTTAGCTAGCTGGTGGTGGTGAAGATCTTG 362
```

```
QY 811 CTTCTTAGAGATCGTGACTCGGGCCCCAAGGTCAACGGCAATGCGGCCTGAGCCTCGCG 870
Db 363 GTCATCCACGATGAACAGAGGGCGCGAAGTGACCTCCAAATGCTGCCCTCACTCTCGGG 422
QY 871 AACTTCTGTGCTGGCAGAGAAGCTGAACAAAGTGAAGTGAACAGCACCAGGAGTACG 930
Db 423 AACTTTTGAATGGCAGAGCAGCAACACCCAGCTGACCGGATGAGAGACATAT 482
QY 931 GACACTGCATCTCTTCAACAGCAGGACGCTGTGGAGCAGCAGCTGTGACACATG 990
Db 483 GACACAGCAATCTTTTCCACAGACAGGACTGTGGTGGTCCAGACATGTGATCTT 542
QY 991 GGCATGCTGTATGGGTACCATGTGTACCCCAAGAAAGTCTCTGTCAATTTGATG 1050
Db 543 GGGATGCTGATTTGGAACCTGTGTGATCCGAGCAGAAAGTCTCGCTCATAGAAGAT 602
QY 1051 GATGGCTTCCATCAGCTTCAACACTGCCACAGCTGGGCCACGCTGTTCAACATGCC 1110
Db 603 GATGGTTTACAGCTGCCCTTCAACACAGCCATGAATTAGGCCACGTTGTTAATCATG 662
QY 1111 CATGACAATGTGAAGTCTGTGAGAGGTGTTGGGAAGCTTCCGAGCACAACCATGATG 1170
Db 663 CATGATGATCAAGCAGCTGTGCCAGCTTAAATGTTGTAACAGGATTTCCACATGATG 722
QY 1171 TCCCGACCCCTCATCCAGATCGACCGTGCACCCCTGCTGGTGAAGCTGTGAGTGTG 1230
Db 723 GGTCAATGCTTTCCAACTGGACACAGCAGCGCTTGGTCTCTTGCAGTGCCTAGATG 782
QY 1231 ATCCAGGACTTCTGGACAGCGGCAGCTGCTGCTCTGTCCTGACCAACCCAGCAAGCC 1290
Db 783 ATACATCATTTCTGGATAATGTTGATGGGAATGTTGATGGACAGCCCTCAAAATCCC 842
QY 1291 ATCTCCCTGCCGAGGATCTCCGGCGCCAGCTACACCTTGAGCCAGAGTGCAGAGTG 1350
Db 843 ATACAGCTCCCGAGGGGATCTCCCTGGCAGCTCGTACGATGCCAACCGSCAGTCCAGTT 902
QY 1351 GCTTTTGGGTGGGTCCAAAGCCCTGCTTACATGCAAGTAC ---TGCACCAAGCTGTGG 1407
Db 903 ACATTTGGGAGGACTCCAAACACTGCCCCGATCGACCGCAGCATGTAGACCTTTGTTGG 962
QY 1408 TGCACCGGAAGGCCAAGGACAGATGCTGCCAGACCCGCCACTTCCCTGTCGGCGGAT 1467
Db 963 TGTACCGGACCTCTGGTGGGTGCTGTGTGTCAAACCAACACTTCCCGTGGCGGAT 1022
QY 1468 GGCACCACTGTGGCAGGAGCAAGCTCTGCTCAAAGGGGCTGGGTGGAGAGACAAAC 1527
Db 1023 GGCACCACTGTGGAGAAAGGAAATGGTGTATCAACGGCAAGTGTGTGAACAAACGAC 1082
QY 1528 CTCACAGCACAGG ---GTGGATGTTCTTGGGCCAATGGGATCCCTATGSCCCC 1581
Db 1083 AGAAAGCATTTTGTATACGCTTTTTCATGGAAGCTGGGAATGTGGGGCCCTTGGGAGAC 1142
QY 1582 TCTCGCGCACATGTGTGGGGGCTGCAGCTGGCGCAGGAGCAGTGCACCAACCCACC 1641
Db 1143 TGTTCGAGAACCTGGGTGGAGAGTCCAGTACACGATGAGGAATGTGAACACCCAGTC 1202
QY 1642 COTGCCAACGGGGCAAGTACTGCGAGGAGTGAAGGTGAATACCGATCTCGAATCTG 1701
Db 1203 CCAAGAAATGGAGGAAGTACTGTGAAGCAAAACGAGTCGCTACAGATCTCTGTAACTT 1262
QY 1702 GAGCCTGCCCCAGCTCAGCTCCCGAAGAGTTCGCGGAGGAGCAGTGTGAGGCTTTC 1761
Db 1263 GAGGACTGTCCAGAC ---AATATGGAAGAAACCTTTAGAGAGGAACAATGTGAAGCAC 1319
QY 1762 AAGCGCTACAACACACAGCACAACCGGCTCACTCTCGCGGTGGCATGGGTGCCCAAGTAC 1821
Db 1320 AAGGATTTTCAAAAGCTTCTTTGGAGTGGCCCTGCGGTGAATGATTTCCCAAGTAC 1379
QY 1822 TCCGGCGTGTCTCCCGGAGCAAGTGAAGTCAATCTCTCGGAGCCAAATGGACATGGCTAC 1881
Db 1380 GCTGGCGTCTCAACAAAGGACAGTGCAGCTCATCTGCCAAGCCAAAGGCAATGGCTAC 1439
```

```
QY 1882 TTCTATGTCTGGCAGCCCAAGGTGGTGACGGCACGGCTGTCCTCTCTGACTCCACCTCC 1941
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1440 TTCTTCGTTTTCAGCCCCAAGGTTCTAGATGTTACTCCATGTCAGCCAGATTCCACCTCT 1499

QY 1942 GTCTGTGTCCTCAAGCAAGTGCATCAAGCTGGCTGTGATGGGAACCTGGGCTCCAAAG 2001
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1500 GTCTGTGTGCAAGGACAGTGTGTAAGAGCTGGTTGTGATCGCATATAGACTCCAAAAG 1559

QY 2002 AGATTGCAACAAGTGTGGGGTGTGGGGAGACAATAAGAGCTCAAGAGGTGACTGGA 2061
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1560 AAGTTTGATAAATGTGGTGTGGGGGGAATGGATCTACTTGTAAAAAATATCAGGA 1619

QY 2062 CTCCTTCCACCAAGCCCATGATGGGTACAAATTCGTGTGGTGGCCATCCCGCAGGCGCTCA 2121
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1620 TCAGTTACTAGTGAACCACTGGATATCATGATATCATCAAAATTCCAATTCGAGCCAC 1679

QY 2122 AGCATCGACATCCGCCAGCGGGTTACAAAGGCTGATCGGGAT-GACAACATACCTGGC 2180
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1680 AACATCGAAGTGAACACGGGAACAGAGGGATCCAGGAACAATGGCAGCTTCTTTCG 1739

QY 2181 TCTGAAGAACAGCCAAAGCAAGTACCTGCTCAACGGGCAATTCGTGTGTGTCGGGGTGA 2240
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1740 CATCAAGCTCTGATGSCACATATATCTTAATGTGTGACTACACTTTGTCACCTTAGA 1799

QY 2241 CGGGACCTGTGTGAGGCGACTGTCTGCGGTACAGCGGCACGGGCACAGCGGTGA 2300
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1800 GCAAGACATTATGACAAAGGTGTGTTTGAGGTACAGCGGCTCTCTCGGCGCATGGA 1859

QY 2301 GAGCTGCGAGCTTCCCGGCCCATCTGGAGCGCTGACCGTGGAGTCTCTCCGTGGG 2360
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1860 AAGATTCGAGCTTTAGCCCTCTCAAGAGCCCTTGACCATCCAGGTTCTTACTGTGGG 1919

QY 2361 GAAGATGACACCGCCCGGGTCCGCTACTCTTCT 2395
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1920 CAATGGCCTTCGACCTAAATTAATACACTACT 1954

RESULT 4
US-09-392-184-7/c
; Sequence 7, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55
; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 3126
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3126)
; OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
US-09-392-184-7

Query Match 21.0%; Score 598.2; DB 4; Length 3126;
Best Local Similarity 57.4%; Pred. No. 2.4e-125;
Matches 1303; Conservative 0; Mismatches 908; Indels 60; Gaps 10;

QY 157 CTCATTTTTCAGATCAGCAGATTCAGGAGGACTTTTACTACACCTGACCGCGGTGCT 216
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3122 CTGCGCTCCACCTGTCGCGCTTCGGCAAGGGCTTCGTGTGCGCTGG-GGCCCGACGAC 3063

QY 217 CAGTCTTGTGCTCCGCGCTTCTCCACTGAGCATCTGGGGGTCCCGCTCCAGGGGCTCACG 276
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3062 AGCTTCTGCGCGCGCACTTCAAGATCGAGCGCTCGGGGGCTCCGCGGGGCGG---ACC 3006
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 277 GGGGGCTCTTCAGACCTGCGACGCTGCTTCTATCTGGGGACGTGAACG-CGAGCGCGGAC 336
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 3005 GGGGGCAGCGGGGGTGCAGCGCTCTCTCTCCGGGACCGCTCACCGCTCCGCGCGAG 2946
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 TCCTTCCGCTGCTGTGAGCCTGTGCGGGGGGCTCCCGGAGCCCTTTGGCTACCGAGCGCGC 396
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2945 TCCTTGGCGCGCTGACGCTGTGCCGCGGCTGAGCGGCTCTCTCTGCTGAGAGCGGAG 2886
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 GAGTATGTCAATAGCCCGCTGCCAATGCTAGGCGCGCGCGCGCGAGCGCAACACCGAG 456
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2885 GAGTTCACCATCCAGCGCAGGCGCGGGGGGCTCCCTGGGTAGCGCGACCGGCTTCAG 2826
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 GCGCGCACACCTTCTCCAGCGCGGGGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCT 516
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2825 CGCTGGGGTCCCGCGGAGCCGCCCTCCCGCGAGAGCCCGAGTGGGAGTGGAGCG 2766
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 CGCTGCGGGGTG-GCCTCGGGGTGGAAACCCCGCCATCTACGGGGCCCTGGACCTTACAA 575
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2765 GGAGAGGGTTCAGAGCGAGGAGAGAGAGACCCAGGAGGACAGCGAGGAGAGAGCCAA 2706
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 576 GCGCGCGCGGGGGCTTCGGGGAGAGTCGTAGCCGGCGCAGGTCTGGGGCGG-G-CC 630
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2705 GAAGAGGAGCAGAAGCGGCTAGCGAGCGCCACCGCCCTGGGGGCCACGAGTAGGACC 2646
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 631 AAGCGTTTCTGTCTATCCCGGGTACGTGGAGACGCTGGTGGTCCGGACAGTCAATG 690
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2645 AAGCGGTTTGTGTCTGAGCGGCTTTCGTGGAGACGCTGCTGTGGCCGATCGTCCATG 2586
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 691 GTCAGATTTCACGGCGCGGACCTGGAAACATTTCTGCTGACGCTGCTGCAACAGCGGGG 750
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2585 GCTGCTTCTACGGGGCGGACCTGTCAGAAACACATCTGACGTTTATGCTGTGGCAGCC 2526
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 751 CGACTCTACCGCCATCCCGAGCATCTCAACCCCATCAACATCGTTGTGCTCAAGTGTG 810
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 CGATCTACAGCACCCCGCATCAAGAATTCATCAACCTGATGGTGTAAAGGTGCTG 2466
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 811 CTTCTTAGAGATCGTGACTCCGGGGCCCAAGGTACCGGCAATCCGGCCCTGACGTGCGC 870
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2465 ATCGTAGAAGATGAAAAATGGGCCCCAGAGGTGTCGACAAATGGGGGCTTACACTGCT 2406
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 871 AACTTCTGTGCTGCGCAGAGAAGAGCTGACAAAGTGTGAGTGTGACAGACCCCGAGTATG 930
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2405 AACTTCTCAACTGGCAGCGCGCTTTCAACAGCCCGACGCGCCGCCACCCAGAGACATAC 2346
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 931 GACACTGCCATCTCTTACACAGGAGGACCTGTGTGG---AGCCACACAGCTGTGACACC 987
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2345 GACACGGCCATCTGCTCACAGACAGAACTTCTGTGGCAGAGGGGCTGTGTGACACC 2286
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 CTGGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAAGTGTCTCTCATTTAG 1047
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2285 CTGGGTGTGGCAGACATCGGACCATTTGTGACCCCAACAAAGCTGCTCCGTGATCGAG 2226
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1048 GAGGATGGGCTTCCATCAGCCTTACCACCTGCCAGAGAGTGGGCCAGCTGTTCAACATG 1107
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2225 GATGAGGGCTCCAGCGGCCCCACACCTTGGCCCATGAATAGGGCAGCTCTCTCAGCATG 2166
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1108 CCCCATGACAATGTGAAAGTCTGTGAGGAGGTGTTGGGAAGCTCCGAGCCCAACACATG 1167
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2165 CCCCAGAGGACTCCAAAGCCCTGCACAGGCTCTTCGGGGCCCATGGGAAGCACCACCTG 2106
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1168 ATGTCCCGACCTTCATCCAGATCGACCTGCCAACCCCTGGTCAAGCTGAGTGTGCC 1227
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2105 ATGGCACCGCTTCTGCTCCACCTGAACCCAGAGGCTGCTGTCCCTCCCTCGACGCGATG 2046
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1228 ATCATCAGGACTTCTTGACAGCGGGCAGCGTGAAGTCTGCTGCTGGACACCCAGCAAG 1287
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2045 TATCTCAGAGAGCTTCTTTGAGGGGGGAC-----CGATTTCAGATAT 2003
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1288 CCCCATCTCCCTGCCGAGGATCTGCCGGGGCGCAGCTACACCTGAGCGCAGAGTGTGAG 1347
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2002 TTATGCGCCTGTGCTAGTGTAGGTGTAGAGGTGTCATTT-----ATGCTCTTTGG 1955
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1348 CTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCTTACATGACGACTGTGACCAAGCTGTGG 1407
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
Db 1954 CCGGATTTCGCCCACTGCCCCAACACCTCTGCTCAGGAGCTCTCGCCCCAGCTTTGGTGC 1895
Qy 1408 TGCACCGGGAGGCCAAGGACACAGATGGTGTGCCAGCCGCCACTTCCCTGGGCCGAT 1467
Db 1894 CACACTGATGGGGTGAGCCCTGTGCCACAGAAATGGCAGCCTGCCCTGGGGTGAC 1835
Qy 1468 GGACCAAGCTGTGGGAGGCAAGCTCTGCTCTCAAGGGCCCTGC-----GTGGAGAG 1520
Db 1834 GGCAGCCGTGGGCCCTGGCCACTCTGCTCAGAGGCAAGCTGTCTACCTGAGGAGA 1775
Qy 1521 ACACAACCTCAAGAACAGCAGAGGTGGATGTTCTTGGGCCAAATGGATCCCTATGGCCC 1580
Db 1774 AGTGGAGAGCCCAAGCCCGTGTAGATGAGGCTGGGCACCGTGGGACCCCTGGGGAGA 1715
Qy 1581 CTGCTCCGCACATGTGGTGGGCGGTGCGAGCTGGCCAGGAGCAGTGCACCAACCCAC 1640
Db 1714 ATGTTCTCGGACCTGTGGAGGAGGTACAGTTTTCACACCGTGTAGTGAAGACCCGA 1655
Qy 1641 CCTTGCACAAGGGGGCAAGTACTCGGAGGAGTGAGGGTGAATACCGATCCTGTCAATCT 1700
Db 1654 GCCTCAAGATGGAGGAAGATACCTGCTGGGTGCGAGAGCCAACTACCACTGATGCCAC 1595
Qy 1701 GGAGCCCTGCCCCAGCTCAGCTTCGGGAAAGAGCTTCCGGGAGGAGCAGTGTAGGCTTT 1760
Db 1594 GGAGGAATGCC-----CCTGACGGGAAAGCTTCAGGGAGCAGCAGTGTGAGAAGTA 1541
Qy 1761 CAACGGCTACAAACACAGCAGCAACCGCTCACTCTCGCCGTGGCATGGGTGCCCAAGTA 1820
Db 1540 TAATGCTCTACAAATACACTGACATGGAGCGGAATCT---CCTGAGTGGGTCCCAAGTA 1484
Qy 1821 CTCGGCGGTGTCTCCCGGACAAAGTCAAGCTCACTGCTCGGAGGCAAGTGGCTACTGGTA 1880
Db 1483 TGTCTGGGTGTCTCCCGGAGCCGCTGCAAGTTGTTCTGCGGAGCCCGGGGAGAGCGA 1424
Qy 1881 CTTCTATGTCTGACCCACCAAGGTGTGGAGCGACGCTGTCTCTCTCACTCCACCTC 1940
Db 1423 GTTCAAGTGTTCGAGGCCAAGGTGATGATGGCACCTGTGTGGGCCAGAAACACTGGC 1364
Qy 1941 CGTCTGTCTCAAGGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGGGGTCTCAAGAA 2000
Db 1363 CATCTGTCTCGTGCCAGTGTGTCAAGGCGGCTGTGACCATGTGTGAGTCTGCTCG 1304
Qy 2001 GAGATTTCGAAGTGTGGGTGTGGGGGAGACAAATAGAGCTGCAAGAGTGAAGTACTGG 2060
Db 1303 GAAGCTGGACAAATGCGGGGTGTGTGGGGCAAGGCAACTCTGTGAGGAAGTCTCCGG 1244
Qy 2061 ACTCTTCAACCAAGCCCATGATGCTACAATTTGCTGGTGGCCATCCCGCAGGCGCCTC 2120
Db 1243 GTCCCTACCCCAACCAATATGCTACATGACATTTGCACCATCCCACTGGTGCCAC 1184
Qy 2121 AAGCATGACATCCGCGAGCGGTTTACAAGGGCTGATCGGGATGACAACTACCTGGC 2180
Db 1183 TAATATTGACCTGATGGCAGTACCTGCTCAAGCGCAACTGCGCAATCTCTGCCATAGA 1124
Qy 2181 TCTGAAGAACCAAGGCAAGTACCTGCTCAAGGGCAATTTCTGTGTGTCTGCGTGTGA 2240
Db 1123 GCTGAAGACGCTGATGGCAGTACCTGCTCAAGCGCAACTGCGCAATCTCTGCCATAGA 1064
Qy 2241 GCGGACCTGTGTGAAGGCACTGTCTGCGGTACAGGGGACGCGGACAGCGGTGGA 2300
Db 1063 GCAGGACATCTGTGAAGGGGACCATCTCTGAAGTACAGCGGCTCCATCCCACTGGA 1004
Qy 2301 GAGCCTGACGCTTCCGCGGCCATCTCTGGAGCCGCTGACCGTGGAGTCTCTCCGT--- 2357
Db 1003 GCGCTGACAGCTTCCGGCCCTTGGCAGAGCCCTCTGACAGTGCAGCTCTGACAGTCCC 944
Qy 2358 GGGGAAGATGACACGCGCCCGGGTCCGCTACTCCTCTTCTATCTGCCCAAGA 2408
Db 943 TGGCGAGGTCTTCCCGCCCAAAAGTCAAAATACACCTCTTTTGTTCCTAATGA 893
```

RESULT 5
US-09-369-364A-8

```
; Sequence 8, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hirskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver: 2.1
; SEQ ID NO 8
; LENGTH: 3638
; TYPE: DNA
; ORGANISM: Mus musculus ADAMTS-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (278)..(2992)
; NAME/KEY: misc_feature
; LOCATION: (3636)
; OTHER INFORMATION: n = T
US-09-369-364A-8
```

Query Match 20.7%; Score 590.4; DB 4; Length 3638;
Best Local Similarity 57.3%; Pred. No. 1.4e-123;
Matches 1334; Conservative 0; Mismatches 906; Indels 90; Gaps 11;

```
Qy 157 CTCATTTTTCAGATCACAGCATTTTCAGGAGGACTTTTACCTACACCTGACGCCGATGCT 216
Db 440 CTCGGCTTCCACCTGTCCGCTTCGGCCAGGGCTTCGTCTGGCCTGGCCCTGACGCC 499
Qy 217 CAGTTCTTTGGTCCGCCCTTCTCCACTGAGCATGTGGCGTCCGCCCTCCAGGGCTCAC 276
Db 500 AGCTTCTTGGCGCGCAATTCAGATCGAGCGCTCGGGGGCTCGAGCGCGCG---GCC 556
Qy 277 GGGGCTCTTACAGACTCGGAGCTGCTTCTTCTGGGAGCTGAACCGCGAGCGCGAC 336
Db 557 GGGGCGAGCCGGGACTCGGTGGCTTCTTCTTCTGACAGTGAATGGAGACGGGAG 616
Qy 337 TCGTTCGTGTGTGAGCTGTGCGGGGGGTCTCCCGGAGGCTTTTGGCTACCGAGGCC 396
Db 617 TCGTGGCGGCGATGAGTGTGTCGCGGCTGGAGCGCTCGTTCTTCTGGCAGCGAG 676
Qy 397 GAGTATGCTATAGCCGCTGCCCAATGTAGCGCGCGCGGCGCGCACCGCAACAGCCAG 456
Db 677 GAGTTTCAACCATCAGCCACAGGGCGCTGGGAGTCCCTGGACAGCCCTCATCGCCTG 736
Qy 457 GCGCACACCTTCTCAGCGCGCGGGGTGTTCGGGGCGGGCTTCCGGGAGA---CCCCAC 512
Db 737 CCGTGGGGCGGGAGACGGCGCGCAAGACCCCGGCTCGCTGCGCGCAAGTTTCCCG 796
Qy 513 CTCTCGCTCGGGGTGGCTCGGGGTGGAAACCCGCCCATCTCTACGGGCCCTGGAGCCCTTA 572
Db 797 CTCCCTCAAGGACTGAGTGGGAGGTGGATGGTAAATGGGCGAGGACAGAGAGAAAGT 856
Qy 573 CAAGCGCGCGGGGGCTTCGGGGAGAGTGTAGCGGCGCGGCTCTGGG----- 624
Db 857 GACAAGAGAGGACAAGAACGAGGAGGGGTGTCTCAAGAGACAGAAAGACTCC 916
Qy 625 -----CGCGCAAGCGTTTCTGTCTATCCCG 651
Db 917 CGCAAGTCCCAACCCCTTCGGATCCAAACTAGAACAGAGAGTTTGTGTCGAGGCT 976
Qy 652 CGGTAGTGGAGCGCTGGTGTGCGGAGCAGAGTCAATGGTCAATTTCCACGGCCCGGAC 711
Db 977 CGTCTGTGAAACACTTCTGGTGGTGTATGCTCATGGTGTGCTTATGGGACCGAC 1036
Qy 712 CTGGACATATCTGCTGACGCTGTGCAACGGCGCGGCTCTACGGCATCCCGAC 771
Db 1037 CTGCAAGAACCATCTCTACGCTGATGTAATGGGAGCGCGCAATTAAGACCCCGAGC 1096
Qy 772 ATCCTCAACCCCATCAACATCGTTGTGGTCAAGGTGCTCTTTAGAGATCGTGACTCC 831
```



```
Db 1097 ATCAGGAACCTCGTCAACCTTGTGTGGTGAAGTCTAATAGTGGAAAAAGAAAGATGG 1156
QY 832 GGGCCCAAGGTCACGGCAANTCGGCCCTGAGCTGCGCAACTTGTGTGCTGGCAGAG 891
Db 1157 GGGCCGGAGTGTCCGACAAACGGGGGCTCACACTGCGCAACTTCTGAGCTGGCAACGG 1216
QY 892 AAGCTGAACAAAGTGAAGTCAACAGCACCCGAGTACTGGGACACTGCCATCTCTTACC 951
Db 1217 CGTTTCAACAGCCGACGACCGCCAGCCGAGCAGTATGACACTGCGCATCTGTTCACC 1276
QY 952 AGGAGGAGCTGTGTGG---AGCCACCACTGTGACACCTGGGCGATGGCTGATGTGGGT 1008
Db 1277 AGACAGAACTTCTGTGGAGGAGGAGCAGTGTGACACCTTGGGATGGCAGAGCTTGGC 1336
QY 1009 ACATGTGTGACCCCAAGAGAAGTGTCTGTCAATGTAGAGCAGTGGCTTCCATCAGCC 1068
Db 1337 ACCATCTGTGACCCGACAAAGAGTGTCTCAGTGATCAAGGATGAGGGACTCCAGCAGCC 1396
QY 1069 TTCACCACTGCCACAGAGTGGGCCACGTGTTCACATGTCCCATGACCAATGTGAAGTC 1128
Db 1397 TACACCCCTGGCCCATGAGCTAGGGACAGTCTTCAGCATGGCCCATGATGATCTAAGCC 1456
QY 1129 TGTGAGGAGGTGTTTGGGAAGCTCCGAGCCAAACACATGATGTCGCCGACCTCATCCAG 1188
Db 1457 TGTGTGAGATGTTTGGGCCCATGGGCAAGTACCACATGATGGGCCATCTTCATCCAC 1516
QY 1189 ATCGACCTGTGCAACCCCTGTGTCAGCTGTGCTGATGAGCAGTATCACCAGTCTCTTGAC 1248
Db 1517 GTGAACAAGAGCGTCCCTGTGCTCCCTGAGTGTCTACCTCACAGAGCTCTCTGGAT 1576
QY 1249 AGCGGGCAGGTGACTGCTCTTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGGAT 1308
Db 1577 GATGCTCAGGAGATGTCTTCTGATGCCGCCACCTCGGTTCTGCCCTCCCAAGGC 1636
QY 1309 CTCGGGGCGGCAG-----CTACACCTGAGCCAGCAGTGCAGCTGGCTTTGGCGTG 1362
Db 1637 CTCGGGGCACAGCACCTCTACGAGCTGGACGAGTGAAGCAGATCTTTGGGGCCT 1696
QY 1363 GGCTCCAAAGCCTGTCTCTTACA-TGCAGTACTGCACCAAGTGTGTGTCACCGGGAAGGC 1421
Db 1697 GATTTCGGACACTGCCCAACACCTCTGTGGAGGACATCTGTGTCCAGCTGTGTGCCGT 1756
QY 1422 CAAGGGACAGATGTTGTC-----AGACCCGCCACTTCCCTGGCCGCGAT 1467
Db 1757 CATCGGATGATGATGAGCCCATTTGOCACAAAGAAATGTAGCCTGCTCTGGCTGAT 1816
QY 1468 GGCACAGCTGTGGCGAGGGCAAGCTCTGCTCAAAGGGGCTCGCGTGGAGAGACACAAC 1527
Db 1817 GGTACACCTGTGGCCCTGGCACCTGTGCTGTGATGGTAGCTGTGTACTCAAGGAGAT 1876
QY 1528 CTCACAAAGCACA-----GGGTGGATGGTTCCTGGGCCAAATGGGATCCCTATGGCCCC 1581
Db 1877 GTGGAGATCCCAAGGCTGTGTAGATGGAGACTGGGTCTCTGGAACCTGGGGACAA 1936
QY 1582 TGCTCGCCACATGTGTGGGGCGTGCAGCTGCCAGGAGGAGTGCACCAACCCCAAC 1641
Db 1937 TGTCTCCACCTGTGTGGAGGATACAATTTCTGAACCGTGAATGTGATAATCCAAATG 1996
QY 1642 CCTGCCAACGGGGCAAGTACTGTGAGGAGTGAAGGTGAATACCGATCTCTCAATCTG 1701
Db 1997 CCTCAGATGGAGGAAGTTTGTCTGGTGAAGAGTCAAGTACCAATATGCAACACA 2056
QY 1702 GAGCCCTGCCCCAGCTCAGCCTCGGAAAGAGCTTCCGGAGGAGCAGTGTGAGGCTTTC 1761
Db 2057 GAGGAATGTC-----CACCAAGCGAAAGAGCTTCCGGGAGCAGCAGTGTGAGAAATAT 2110
QY 1762 AACGGCTACACCAACAGCAGCAGCAGCTCACTCTCCCTGGCATGGGTGGCCAAAGTAC 1821
Db 2111 AATGCCACAAACCA---CACTGACCTGGATGGGAATTTCTGCGATGGGTGCCCAAGTAT 2167
QY 1822 TCCGGGCTGTCTCCCGGGACAGTGCAGCTCATCTGCCGAGCCCAATGSCACTGGCTAC 1881
```

```
Db 2168 TCAGGACTGTCCCCCGGAGACCGATGCAAGCTGTTTTCAGAGCCCGTGGGAGGATGAG 2227
QY 1882 TTCTATGTGTGGCACCAAGGTGGTGGACGGCAGCTGTGCTCTCTGACTCCATCTCC 1941
Db 2228 TTCAAAGTGTGTAAGTAAAGTATGATGATGGCACTCTGTGTGGACCGGATACTCTGTC 2287
QY 1942 GTCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGATGGGAACCTTGGCTCCAAAG 2001
Db 2288 ATCTGCTCGGGGGCAATGTGTAAGCTGGCTGTGACCATGTGGTGAACCTCACCTAAG 2347
QY 2002 AGATTCCACAAAGTGTGGGTGTGTGGGGGAGACAATAAGAGCTGCAGAGAGTGAATGA 2061
Db 2348 AAGCTGACAAATGTGGGGTGTGTGGGGGCAAAAGGCATCCCTGTAGGAAGATCTCCGT 2407
QY 2062 CTCTTCAACCAAGCCATGCTATGCTACAATTTCTGTGTGGCCATCCCGCAGCGCTCA 2121
Db 2408 TCITTCACCCCTTCAGTTATGCTACATGACATTTGTACCATCCAGCTGGTGGCACA 2467
QY 2122 AGCATGACATTCGCCAGCGGTTTACAAAGGCTGATCGGGGATGACAACTACCTTCT 2181
Db 2468 AACATTGATGTGAACAGCGGAGTCAACCCAGGGTCAAGGAACGACGCGACTACCTGGCG 2527
QY 2182 CTGAGACACAGCCAAAGCAAGTACCTGCTCAACGGGCAATTTCTGTGTGGCGGTGGAG 2241
Db 2528 CTGAAGACAGCCAAATGGCGAGTACCTGCTCAATGTAACCTGGCCATCTCTGCCATAG 2587
QY 2242 CGGCACTGTGTGAAGGCACTGTGCTGCTACAGCGGCACGCGCAGCGTGGAG 2301
Db 2588 CAAGACATCTGTGTGAAGGGACCATCTCTGAAGTACAGTGGCTCCATGGGTACCTTGGAG 2647
QY 2302 AGCCTGAGCGTTCGCCGGCCATCTCTGGAGCCCTCAACCGTGGAGTCTCTTCTCTCT 2358
Db 2648 CGGCTGAGAGCTTCCAGGCCCTGCTGAGCCCTTTACAGTACAGCTCTCTGACTGTCT 2707
...QY 2359 GGAAGATGACACGCCCGCGGTCCGCTACTCTCTTCTATCTGCCCAAGA 2408
Db 2708 GGTGAGGTCTTCCTCCAAAGTCAGATATACCTTCTTTGTCCTCAATGA 2757

RESULT 6
US-09-122-126B-1
; Sequence 1, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (406)..(2916)
US-09-122-126B-1

Query Match 20.4%; Score 583.2; DB 4; Length 4192;
Best Local Similarity 60.4%; Pred. No. 6e-122;
Matches 1093; Conservative 0; Mismatches 693; Indels 24; Gaps 7;
```


Db 944 CTGACGCTCCATGGCGGGTGTATGCGGGGCTCGAGCATTAACCTGTGACCCCTGG 1003
QY 737 TGGCAACGGGGGCGGACCTCTACCGGCATCCAGCATCTCAACCCCATCAACATCGTTG 796
Db 1004 CTCCATCCGCAATAGGCTGTACAGCCATGCTAGCATGAGAACACACATCCGCTGGCGG 1063
QY 797 TGGTCAAGTGTGCTTCTTAGAGATCGTGAATCCGGGCGCCCAAGCTCACCGCAATCGG 856
Db 1064 TGGTGAAGTGTGCTTAGCGGACAGGACAAGAGCGCTGGAAGTGAGCAAGAACGCTG 1123
QY 857 CCTGACGCTGGCGCACTTCTGTGCTGSCAGAAGAAGTGAACAAAGTGAAGTGACAAGC 916
Db 1124 CCACCACTCAAGAAGCTTTTCAAGTGGCAGCACCAACACACAGCTGGAGATGACC 1183
QY 917 ACCCGAGTACTGGGACACTGCCATCTCTTACAGGAGGAGACACTGTGTGGAGCCACA 976
Db 1184 ATGAGGAGCACTACGATGAGTATCTCTTACTCGGAGGAGTATTTCTGGGCATCAT 1243
QY 977 CCTGTGACACCTGGGCGATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGCT 1036
Db 1244 CATGTGACACCTGGGAATGGCAGAGCTTGGGACCATATGTTCTCCAGAGCGCAGCTGT 1303
QY 1037 CTGTCAATGAGACAGATGGCTTCCATAGCCTTCCACACTGCCACAGCTGGGCCACG 1096
Db 1304 CTGTGATTGAAGACGATGCGCTTCCAGCAGCGCTTCACTGTGGCTCACGAATCGGACAT 1363
QY 1097 TGTTCACATGCCCCATGACATGTAAGTCTGTGAGGAGTGTGGGAGCTCCGAG 1156
Db 1364 TACTTGGCTCTCCCATGACGATTCGAATCTGTGAAGAGACCTTTGGTTCACAGAG 1423
QY 1157 CCAACCATGATGTCCCGACCTTCATCCAGATCGACCGTCCCAACCCCTGGTCAGCCT 1216
Db 1424 ATAAGCGCTTAATGTCTTCCATCTTACCAGATGTATGATCTTAAGCCCTGGTCCAAAT 1493
QY 1217 GCAGTGTCTCCATCATACCGACTTCTTGGACAGGGGCGAGTGTACATCGAGTACCA 1276
Db 1484 GCACCTCAGCCACCATCAGAAATTCCTGGATGATGGCCATGGTAACCTGTGTGGTGGAC 1543
QY 1277 AACCAGCAAGCCCATCTCCCTGCGCGAGGATCTGCGGGGCGGCGAGCTACACCCCTGAGCC 1336
Db 1544 TACCAGGAAGCAGATCTCGGCCCGGAGAACTCCAGGACAGACCTTACCATGCCACCC 1603
QY 1337 AGCATGGAGCTGCTTTTGGGCTGCGCTCCAGCCCTGCTTACATCGAGTACCTGCA 1396
Db 1604 AGCAGTGAACCTGCATCTCGGGCTGAGTACTCGGTGTGTCGGCGCATGATGCTGTG 1663
QY 1397 CCAAGCTGTGTGACCGGGAAGGCAAGGACAGATGGTGTGCGAGACCCGCACTTCC 1456
Db 1664 CTCGCTGTGTGTGCTGTGTGACCCAGGGCCAGATGGTGTGTGACCAAGAGCTGC 1723
QY 1457 CTTGGGCGATGGCACCACTGTGTGCGAGGGCAAGCTGTGCTCAAAGGGGCGCTGCGTGG 1516
Db 1724 CTGCGGTGGAAGGAGCGCTTGTGGAAGGGGAGAACTGCTCGAGGGCAAAATGTGTGG 1783
QY 1517 AGAGACAACTCAACAGCACAGGTG-----GATGGTCTCGGCCCAAAATGGGATC 1570
Db 1784 ACAAACCAAGAAAATATTTATTCAGCTCAAGCCATGGCAACTGGGATCTTTGGGAT 1843
QY 1571 CTTATGCGCTGTGCGGCACATGTGTGGGGGCGCTGACCTGCGCCAGGAGGAGTGA 1630
Db 1844 CTTGGGCGCATGTTCTCGCTCATGTGTGAGGAGGAGTGCAGTTTGCCTATCGTCACTGTA 1903
QY 1631 CCAACCCACCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGAAGGTGAATACCGAT 1690
Db 1904 ATAACCTGTCTCCAGAAACAACGAGCGTACTCCAGGGAAGAGGGCCATCTACCGCT 1963
QY 1691 CCTGCAATCTGGAGCGCTGCCCGAGCTAGCCTCCGGAAGAGCTTCCGGAGAGAGCAGT 1750
Db 1964 CTTGAGTCTATCGCTGCG-----CACCAATGGTAAATCATTTTCGTTCATGAACAGT 2017
QY 1751 GTGAGGCTTTCAACGGCTTACACACACAGCACCAGCCGCTCACTCTCGCGTGGCATGGG 1810

Db 2018 GTGAGGCCAAAATGGCTATCAGTCTGATGCAAAAGAGTGCAAAACTTTTGTGGAATGGG 2077
QY 1811 TGCCCAAGTACTCGGGGTGTCTCCCGGGGCAAGTGCAGAGCTCATCTGCGCAAGT 1870
Db 2078 TTCCAAATATGCAGGTGTCTCTCCAGCGATGTGTGCAAGCTGACCTGCAGAGCCAAAG 2137
QY 1871 GCATGCTGCTACTTATGTGCTGGCACCAAGGTGTGGAGCGGACCGCTGTGCTCTG 1930
Db 2138 GCATGCTGCTACTTATGTGCTGCTTTCTCAAAGGTGACCGATGCAATGTAGGCGGT 2197
QY 1931 ACTCACCTCGTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGG 1990
Db 2198 ACATTAATTCGCTGTGCTCGGGGGAAGTGTGTGAAGTGGCTGTGACCGCATCATG 2257
QY 1991 GCTCAAAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGACAAATAAGAGCTG 2050
Db 2258 GCTCAAAGCTGCAGTATGACAAGTGGCGAGTATGTGGAGGAGACAACCTCCAGCTG 2317
QY 2051 AGGTGACTGGACTCTTCCACCAAGCCCATGATGGCTACAAATTTGCTGGTGGCCATCCCG 2110
Db 2318 AGATTTGTGAACCTTTAATAAGAAAAGTAAGGTTTACACTGAGCTGGTGAAGATTCTG 2377
QY 2111 CAGCGCTCTCAAGCATCGACATCCGCCAGCGGCTTACAAAGGCTGATCGGGGATGACA 2170
Db 2378 AAGGGCAACCCACATAAAGTTCGACAGTTCAAAGCCAAAGACCACTAGATTCACTG 2437
QY 2171 ACTACCTGGCTCTGAAGAACAGCCAGCAAGTACCTGCTCAAGGGGCAATTCGTGGTGT 2230
Db 2438 CCTATTTAGCCCTGAAAGAAACCGGTGAGTACCTTATCAATGGAAGTACATGATCT 2497
QY 2231 CGGCGGTGGAGCGGACCTGTGTGAAGGCGAGTGTGCTGCGGTACAGCGGCACGGCA 2290
Db 2498 CCACCTCAGAGCATATCATTTGACATCAATGGAACAGTCAATGAACTATAGCGGTGGAGCC 2557
QY 2291 CAGCGGTGGAGAGCCCTGCA-----GGCTTCCCGGCCCATCTCTGGAGCCGCTGACCGTGG 2344
Db 2558 ACAGGGATGACTTCTCTCATGCGCATGGCTACTCTGCCAGGAAGAAATCTTAATAGTGC 2617
QY 2345 AGTCTCTCTCGGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCCTTCTATCTGCCCA 2404
Db 2618 AGATTTCTGCAACAGAGCCCACTAAACCAATTAGATGTCGGTTATAGCTTTTGTTCCTCA 2677
QY 2405 AAGAGCC 2411
Db 2678 AGAGTC 2684

RESULT 8

US-09-369-364A-1

; Sequence 1, Application US/09369364A

; Patent No. 6391610

; GENERAL INFORMATION:

; APPLICANT: Apte, Suneel

; APPLICANT: Hurskainen, Tiina L.

; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

; FILE REFERENCE: 26473/4007/10-30-00

; CURRENT APPLICATION NUMBER: US/09/369,364A

; CURRENT FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 3002

; TYPE: DNA

; ORGANISM: mus musculus ADAMTS-5

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (18)..(2810)

US-09-369-364A-1

Query Match

Best Local Similarity 19.2%; Score 547.4; DB 4; Length 3002;

Matches 1043; Conservative 0; Mismatches 746; Indels 18; Gaps 3;

LOCATION: (2)..(2623)
US-09-369-364A-14

Query Match 16.5%; Score 469.8; DB 4; Length 2625;
Best Local Similarity 56.4%; Pred. No. 1.5e-96;
Matches 1008; Conservative 0; Mismatches 752; Indels 27; Gaps 6;
QY 629 CCAAGCGTTTCGTCTATCCCGGTCAGTGGAGACGCTGGTGTGCGGACGAGTCAA 688
DB 372 CCAAAACGCTTCTGTCTACCCACGCTTTGTAGAGGTGATGGTGGTGCACACAGGA 431
QY 689 TGGTCAAGTTCCACGGCGGGAGCTGGACATATCTCTGACGCTGTGTCGCAACGGCGG 748
DB 432 TGGTTTATACCAACGGAGCAACCTTCAACATATATCTTAACCTTAAATCTCCATTGTAG 491
QY 749 CGCGACTCTACCGGCATCCAGACTCTCAACCCCATCAACATCTGTTGTCTCAAGTGC 808
DB 492 CTCTCTCTATAAAGACTCAAGTATGGAAATTTAAATTAATTTATTTATTTGTAACCTTAG 551
QY 809 TGCCTCTTAGAGATCGTGACTCCGGGCCCAAGGTCACCGCAATGCGGCCCTGACGCTGC 868
DB 552 TTGTGATTCATAAATCAACAGGAAGACCTTACATAAATTTCAAAGCCAGACACATTA 611
QY 869 GCAACTCTGTGCTGGCAGAGAGCTGACAAAGAGTGAAGTGAAGCAACCCGAGTACT 928
DB 612 AGAAGCTTTGGCAGTGGCAGCACTCAAAAGAACTACTTGGGTGGATTTCAGCAC----- 664
QY 929 GGGACACTGCCATCTCTTCCAGGAGGACCTGTGTGGAGC---CACCACTGTGACA 985
DB 665 --GACACAGCGTCTGTGTCACAGGGAAGATCTGCAGAGCTCAGGACAAATGTGACA 722
QY 986 CCCTGGGATGGCTGATGTGGTACCATGTGACCCCAAGAGAGTGTCTGTCTATTG 1045
DB 723 CTTTAGGCTTGTGTAAGTGGGAACCATTTGCGACCCCTACCGAAGCTGTTCATTAGTG 782
QY 1046 AGGAGATGGCTTCCATCAGCTTCACTGCTGACCTGCCAGCTGGGCAAGTGTTCACA 1105
DB 793 RAGACATGGCTGAGCACAGCTTTCAATAGCTCAGAGCTGGGCCATGTGTTTAATA 842
QY 1106 TGCCCATGACATGTGAAAGTCTGTGAGGAGGTGTTTGGAAAGCTCCGAGCCCAACACA 1165
DB 843 TGCCCTCAGATGACAGCAATAAATGCAAGGAAGAGGTTAAG---AGTCCCGAGCATG 899
QY 1166 TGATGTCCCGACCTCTATCAGATCGACCTGCCACCCCTGTGAGCTGTGAGCTGTG 1225
DB 900 TCATGTGACCAACACTGAAGTCTTACACCAACCCCTGGATGTGCTCAAGTGCAGTCGGA 959
QY 1226 CCATCATCAGGACTTCTGTGACAGCGGACAGCTGCTGCTGACCAAC---CCA 1282
DB 960 AATACATCAGTGTCTTAGACACTGGTACGAGAGTGTGCTGTAATGAACCTCGAT 1019
QY 1283 GCAAGCCCATCTCCCTGCCCCAGGATCTGCGGGCGCCAGCTACACCTGAGCCAGCAGT 1342
DB 1020 CCAGGACCTATCTTGTGCTTCCCACTGCGCGGCTTCTCTACAAGCTGATAAACAAT 1079
QY 1343 CGAGCTGGCTTTTGGCTGGGCTCCAGCCCTCTCTATCATGCTAGTGTGACCAAGC 1402
DB 1080 GTGAAGTATTTTGGGCCAGGCTCTCAAGTGTGCGCCCTATATGATGAGTGCAGAGCGG 1139
QY 1403 TGTGTGTGACCGGGAAGGCCAAGGACAGATGTTGTCGACAGCCGCACTTCCCTGGG 1462
DB 1140 TCTGGTGCATAAATGTGATGGAGCACACAAAGGCTGCAAGCTGCAAGCTGATAAACAAT 1199
QY 1463 CCGATGGACAGCTGTGGCAGGGCAAGCTCTGCTCTCAAGGGGCTGTGTCGAGAGAC 1522
DB 1200 CAGATGGAACCGAGTGTGAGCTGGAAGCACTGCAAGTTTGGATTGTGTTCCCAAG 1259
QY 1523 ACACCTTCAACACACAGGATGATGTTCTCTGGGCAATGAGTATGATGATGATGATG 1582
DB 1260 AAA---TGGAGGGCCCTGCAATTTGATGATCTCTGGGAGGTTGAGGCACTTTGGGACCT 1316
QY 1583 GCTCGGCACATGTGTTGGGGCTGTGAGCTGCGCAGGAGGAGTGCACCAACCCACCC 1642

DB 1317 GCTCAAGAACGCTGTGGAGGAGGCGATCAAAACAGCCATCAGAGTGCACAGACACG 1376
QY 1643 CTGCCAACGGGGCAAGTACTGGGAGGAGTGGAGGTGAATACCGATCCTGCAATCTGG 1702
DB 1377 CAAAAAATGTTGGGAAGTACTGTGTAGGAAGAGAAATGAAGTTCAAACTCTGCAACACGG 1436
QY 1703 AGCCCTGCCCGCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGCTTTCA 1762
DB 1437 AGCCCTGCATGAAGCAGAAAGCGAG-----ACTTCCGAGGAGCAGTGTGCTACCTTTG 1490
QY 1763 ACAGCTAAACACAGCAGCAACCGGCTCACTCTCGCCGCTGGCATGGTGCCTCCCAAGTACT 1822
DB 1491 ATGSCAAACACTTCAACATCAATGGTCTGCTGCCAGCGTACCTGCTTCTTCTTAAGTACA 1550
QY 1823 CGGCGTGTCTCCCGGGGACAAGTGAAGCTCATCTGCTCCGAGCAATGGCACTGGCTACT 1882
DB 1551 GCGGAATTTTGTATGAAGAGCGGTTGCAAGTTGTTCTGACAGTGGCAGGAACACAGCCT 1610
QY 1883 TCTATGTGCTGGCACCACCAAGTGTGGAGCGCAGCTGTGCTCTCTGACTCCACCTCCG 1942
DB 1611 ACTACAGCTCCGAGCAGAGTGAATGACGGAACCCCTTGTGGCAGGACACAAATGACA 1670
QY 1943 TCTGTGTCCAAGCAAGTGCATCAAGCTGGCTGTGTGATGGAAACCTGGGCTCCCAAGAAGA 2002
DB 1671 TCTGTGTCCAAGGCTTTTGGCGCAAGCTGGATGGATCATATTTTAACTCAAGGTCC 1730
QY 2003 GATTCCGCAAGTGTGGGCTGTGTGGGGGAGACAATAGAGCTGCAAGAGGTGACTGGAC 2062
DB 1731 GGAAGATAAATGTGGGATTTGTGGTGGAGATAATTTCTCATGCAAAACAGTGGCAGGAA 1790
QY 2063 TCTTTCACCAAGCCCATGCTAGCTGCTACAAATTTCTGTGTGGCCATCCCGCAGGCGCTCAA 2122
DB 1791 CATTTAACTGTCTCCATTATGTTTACAATACTGTTTCCGAATTCGGCTGGTGTCTACCA 1850
QY 2123 GCATGACATCCCGCAGCGGCTTACAAAGGCTGTATCGGGGATGACAACTACCTGGCTC 2182
DB 1851 GCATTTGACGTGCTGACGACAGCTTCTCAGGGAAGTCTGAGGATGACAACTACCTAGCTT 1910
QY 2183 TGAAGAACGCCAAGCAAGTACTGCTCAACGGGCAATTTCTGTGTGCTGCGCGGTGGAGC 2242
DB 1911 TATCAACAGTAAGGTGAATTCCTGCTAAATGGAGACATTTGTTCTCCATGTCCAAA 1970
QY 2243 GAGACCTGTGTGAAGGCGAGTCTGCTGCGGTACAGCGCAGCGGCGCAGCGGTGGAGA 2302
DB 1971 GCGAGTCCGCGTGGGAGCGCGTCAATGAGTACAGCGGATCGGACATGTGTGANA 2030
QY 2303 GCTGTGAGGCTTCCCGGCCCATCTCGGCGCGTACCGGTGAGGTGCTCTCCGTGGGGA 2362
DB 2031 GACTGAAGTGTACGGGACCGTATCGAGGAAGAACTTCTCTTCAAGTGTGTCCTCGTGGGAA 2090
QY 2363 AGATGACACCGCCCGGCTCCGCTACTCTCTTCTATCTGCCCCAAAGAG 2409
DB 2091 AGCTGTATAACCCAGATGTGCGGTACTCATTAATATTCCTCATTTAG 2137

RESULT 10

US-09-369-364A-12
; Sequence 12, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 5804
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-9

FEATURE:
 NAME/KEY: CDS
 LOCATION: (3) (5648)
 NAME/KEY: misc_feature
 LOCATION: (1406)
 OTHER INFORMATION: n-T
 NAME/KEY: misc_feature
 LOCATION: (1563)
 OTHER INFORMATION: n-T
 US-09-369-364A-12

Query Match 14.4%; Score 409.6; DB 4; Length 5804;
 Best local Similarity 54.3%; Pred. No. 6.3e-83;
 Matches 970; Conservative 0; Mismatches 786; Indels 30; Gaps 6;

QY 629 CCAAGCGTTTGTCTATCCCGGTTACGTGGAGCGCTGGTGGTGGCGGACGAGTCAA 688
 DB 700 CAAAGCGTTTATCTATCCCGGTTTGTAGAGTCTTGGTGGGACAGACAGAA 759
 QY 689 TGGTCAAGTTCCACGGCGGACCTGGAACATTAATCTGTGACGCTGTGGCAACGGCGG 748
 DB 760 TGGTTTACATGAGGAGAAACCTTCAACACTATATTTTAACTTTAATGTCAATTGTAG 819
 QY 749 CGGACTCTACCGCCATCCAGCATCTCAACCCATCAACATCGTGTGCTCAAGTGC 808
 DB 820 CCTCTATCTATAAGAGCCCAAGTATGGAATTTAATTAATTTAATTTGTGAATTA 879
 QY 809 TGCCTTCTAGAGATCGTGACTCCCGGCGCCAAAGGTCAACCGCAATGCGGCCCTGACCGTGC 868
 DB 880 TTGTGATTCATAAGACAGATGGCCCTCCCATATCTTTAATGTCTGACACATTA 939
 QY 869 GCAACTCTGTGCTGGAGAGAGCTGAACAAAGTGAAGACAGCACCCCGGAGTACT 928
 DB 940 AAAACTTTTGGCAGTGGCAGCATTCGACACAGTCCAGGTGGAATCCATCA 988
 QY 929 GGGACACTGCCATCTCTTACACAGGAGGACCTGTGTGGAGC---CACTACCTGTGACA 985
 DB 989 -TGATCTGTCTCTCTTACACAGAGAGATATCTGACAGCTCAGCAATATGTGATA 1047
 QY 986 CCCTGGGATGGCTGATGGGTACCATGTGTGACCCCAAGAGAGCTGTCTGTCAATG 1045
 DB 1048 CTTTAGGCTGGCTGAATGGGAACCATTTGTGATCCCTATAGAAGCTGTCTATTAGTG 1107
 QY 1046 AGGAGATGGCTTCCATCAGCTTACCCTGCGGAGAGCTGCGGACAGTGTTCACA 1105
 DB 1108 AAGATAGTGATTTAGTACAGCTTTTACGATCGCCCATGAGCTGGGCCATGTGTTAACA 1167
 QY 1106 TGCCCATGACAATGTGAAAGTCTGTGAGGAGTGTGTTGGAGCTCCGAGCCCAACACA 1165
 DB 1168 TGCCTCATGATGACACACAAATGTAAAGAGAGAGTTAG---AGTCCCGCAGCATG 1224
 QY 1166 TGATGTCCCGACCTCATCCAGATCGACCGTGCACCCCTGGTACGCTGAGTGTGCTG 1225
 DB 1225 TCATGGCTCCCAACACTGAATCTTACACCAACCCCTGGATGTGTTGTTAGTGTAGTGA 1284
 QY 1226 CCATCATCAACGACTTCTGTGACAGGCGGACGTGCTGCTCTGTGACCAAC---CA 1282
 DB 1285 AATATATCACTAGTTTTTATGACACTGTGTATGGGAGTGTGCTTAACTGAACCTGAAT 1344
 QY 1283 GCAAGCCCATCTCCCTGCGGAGGATCTGCGGCGGCGGACGTACACCTGAGCCAGCAGT 1342
 DB 1345 CCAGACCCCTACCTTTGGCTGTCCAACTGCCAGGCGCATCTTTACACGTTGAATAACAAT 1404
 QY 1343 GGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTCTCTTACATGACTATGACCAAGC 1402
 DB 1405 GNGAATTTGATTTTGGACAGGTTCTCAGGTGTGCCATATATATGATGAGTGCAGAGCGC 1464
 QY 1403 TGTGTGTCAACCGGAGGAGGAGGAGATGTGTGCGAGACCGGACCTTCCCTCTGGG 1462
 DB 1465 TCTGTGTCAATTAAGCTCAATGGAGTACAAAGGCTGCGGAGCTCAGCACACACCTTGG 1524
 QY 1463 CCGATGGACACAGCTGTGGGAGGAGGAGCTTCCCTCAAGGGGCTGCTGGAGAGAC 1522

DB 1525 CCGATGGACGAGTGGCGAGCCTGGAAAGACTGCAAGNATGGATTTTGTGTCCCAAG 1584
 QY 1523 ACAACCTCAACAAGCACAGGGTGGATGTTCTCTGGGCCAAATGGGATCCCTATGGCCCT 1582
 DB 1585 AAA---TGGATGTCCTCCGTGACAGATGATCTTGGGAAGTTGGAGTCCCTTTGGAACCT 1641
 QY 1583 GTTCGCGACATGTGTGGGGCGTGCAGCTGGCGAGGAGGAGTGCACCAACCCACCC 1642
 DB 1642 GCTCCAGAACATGTGGAGGGGCATCAAAACAGCCATTCGAGAGTGCACAGACAGAAC 1701
 QY 1643 CTGCCCCAGGGGCAAGTACTCGAGGGAGTGAAGGTGAATACCGATCTCTGCAATCTGG 1702
 DB 1702 CAAAATATGTTGGAATATCTGTAGGACGTAGTAATGAATTAAGTCTTGCACACGG 1761
 QY 1703 AGCCCTGCCCGAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGAGTGTGAGGCTTTCA 1762
 DB 1762 AGCCATGCTCAAGCAGAAGCGAG-----ACTTCCGAGATGAACAGTGTCTCACTTTG 1815
 QY 1763 AGGGCTACACACAGCAGCAGCCGCTCACTCTGCCGTGCGATGGGTGCCAAGTACT 1822
 DB 1816 ACGGGAAGCATTTTAACATCAACGCTGTCTCCCAATGTGCGCTGGGTCCCTTAATACA 1875
 QY 1823 CCGGCTGTCTCCCGGAGCAAGTCAAGCTCATCTGCCGAGCCAATGGCACTGG---ACT 1882
 DB 1876 GTGGAATTTGTATGAAGACCGTGCAGTGTCTTCTGAGAGTGCAGGAAACACAGCT 1935
 QY 1883 TCTATGTGCTGACCCCAAGTGTGGAGGACGCTGTCTCTCTGCTGCTGCTGCTGCTGCT 1942
 DB 1936 ACTATCAGCTTCCGAGACAGTGTAGATGGAACCTCTTGTGGCCAGGACACAATGATA 1995
 QY 1943 TCTGTCTCCAAAGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGGCTCCCAAGA 2002
 DB 1996 TCTGTCTCCAGGCGCTTTCGCCGCAAGCTGTGATGCGTCAATGTTTAACTCAAAAGCCC 2055
 QY 2003 GATTGACAAAGTGTGGGTGTGTGGGAGGACAATAAGAGCTGCAAGAAGTGAATGAC 2062
 DB 2056 GGAGAGATAAATCGGGTTTGTGTGGCGATAATTTCTTATGCAAAAACAGTGGCAGAA 2115
 QY 2063 TCTTACCAGCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122
 DB 2116 CATTTAATACAGTACATTTATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2175
 QY 2123 GATCTGACATCCGACGAGCGCTTACAAGGCTGATCGGGATGACAACTACCTGGCTC 2182
 DB 2176 ATATGATGTGGCAGCAGCTTCTCAGGGGAAACAGACGATGACAACTACTTACTT 2235
 QY 2183 TGAAGAACCAAGCAAGTACCTGCTCAACGGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2242
 DB 2236 TATCAAGCAGTAAAGTGAATTTCTTGTAAATGGAACCTTTTGTCTCAATGGCCAAA 2295
 QY 2243 GGGACCTGTGTGAAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2302
 DB 2296 GGGAAATTCGATTTGGGAATGCTGTGGTAGATGATGAGTGGTCCGAGACTGCCGTAGAA 2355
 QY 2303 GCTGAGGCTTCCCGGCGCATCTCCGAGCGCTGACCGTGGAGGCTCTCTCTGCTGCTGCTGCT 2362
 DB 2356 GAATTAACCTCAACAGATCCATTTAGCAAGAACTTTTGTCTCAGGTTTGTGCTGCTGCTGCT 2415
 QY 2363 AGATGACACCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2408
 DB 2416 AGTTGTACAAACCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2461

RESULT 11

US-09-130-491-7

; Sequence 7, Application US/09130491

; Patent No. 6416974

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001

; CURRENT APPLICATION NUMBER: US/09/130,491


```
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2114
; TYPE: DNA
; ORGANISM: Rattus rattus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1445)
US-09-130-491-7

Query Match      11.1%; Score 318; DB 4; Length 2114;
Best Local Similarity 57.8%; Pred. No. 1.8e-62;
Matches 705; Conservative 0; Mismatches 470; Indels 45; Gaps 6;

QY 1204 CCTGGTCTGAGCTGCTGCCATCATACCGACTTCTCTGGACAGGGGCGACGGTGC 1263
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 CCTGGTCTGCTGAGTGTCTTACCTCAGGAGCTCTCTGGATGATGCTCAGGAGAC 80
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1264 TCCCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCGAGAGTCTGCCGGGCCAGC 1323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 TCCCTCTGGATGATGGCCACAGCACCTCTATGAGCTGGACACAGT----- 129
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1324 TACACCTTGAGCCAGCTGGAGTGCTTTTGGCGTGGGCTCCAAAGCCCTGCTTAC 1383
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 -----GCAAGCACATCTTTGGGCTGATTTCCGACACTGCCCAACACCTCTGTG 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1384 ATGCAGTACTGCACCAAGCTGTGTGCACCGGGAAGGCCAAGGACAGATGGTGGCC-- 1441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 GAGGACATCTGTGCCAGCTGTGTGCGTTCATCGGATAGTATGAGCCCATTTGCCAC 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1442 ---AGACCCGCCACTTCCCTGGGCCGATGGCACCAGCTGTGGGAGGCAAGCTTGC 1497
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 ACAAGAATCCAGCTGTCTGGGCTGATGTACGCCCTGTGGCCCTGGGCACCTGTGC 299
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1498 CTCGAAGGGGCTGTGTGGAGAGACACAACCTCAACAGCAC-----GGTGGATGGT 1551
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 CTGGATGTAGCTGTGTCTCGGAGGAAGTAGAGATCCCAAGGCTGTGTAGATGGA 359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1552 TCTTGGGCCAAATGGATTCCTATGTGCCCTGTCTGCGCACATGTGGTGGGGCGTGCAG 1611
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 GACTGGGTCTCTGGGACCTGGGACAAATGTTCTCGCACCTGTGTGGAGGATACAG 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1612 CTGGCCAGGAGCAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGGAGGGA 1671
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 TTTTCGAACCTGTAGTGTGAATCCAGACCTCAGAAATGGAGAAATTTTGCTGGGA 479
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1672 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAAG 1731
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GAGAGATCAAGTACCAATCTTGCACAGACAGAGAAATGTC-----CACCAATGGAAA 533
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1732 AGCTTCGGGAGGAGCAGTGTAGGCTTTCAACGGCTTACAAACACACAGCAACCGGCTC 1791
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 AGCTTCAGGAGGAGCAGTGTGAATAATATATGCTTACAACCA---CAGGGACCTGGAT 590
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1792 ACTCTGCGCGTGGATGGTGGCCCAAGTACTCGGGGTGTCTCCCGGGGCAAGTGCAG 1851
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 591 GGAATTTCTCTAGTGGTCCCAATAATCTCAGGAGTGTCCCGCCGAGACCGATGCAAA 650
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1852 CTCATCTGCCGAGCAATGCACTGCTACTTCTATCTGTGGCACCCCAAGGAGTGGTGCAC 1911
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 651 CTGTTTTCAGAGGCCCTGGGAGAGTGTCAAGTGTGTTGAAACTAAGGTGATCGAT 710
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1912 GGCAGCTGTCTCTCTGACTCCACCTCGCTGTGTCTCAAGGCAAGTGCATCAAGGCT 1971
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 711 GGCAGCTGTGTGGACCGGATCTCTGCGCATCTGTGTGGGGGACAGTCCGTAAAGCT 770
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1972 GGCTGTATGGGAACCTGGGCTCCAAAGAGATTCGACAAGTGTGGGGTGTGTGGGGGA 2031
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 12

```
US-09-369-364A-10
; Sequence 10, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(737)
US-09-369-364A-10

Query Match      9.1%; Score 259.4; DB 4; Length 739;
Best Local Similarity 64.4%; Pred. No. 1.9e-49;
Matches 405; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 629 CCAAGCGTTTCGTGTCTATCCCGCGGTAGCTGGAGACGCTGGTGGCGGACGAGTCAA 688
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55 CCAAGCGTTTCGTGTCTATGAGCGCGCTTCGTGGAGACGCTGGTGGCGGATGCTCA 114
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 689 TGGTCAAGTTCACGGCGGGACCTCGGACATTTATCTGTGACGCTGTGTGGCAACGGGG 748
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 115 TGGTGCCTTCTACGGGGCGACCTCGACAGACCATCTGACGTTAATGTCTGTGGCAG 174
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 749 CGCGTCTTACCGCCATCCAGCATCTCAACCCCATCAACATCTGTGGTCAAGTGC 808
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 175 CCGGATCTTCAAGCACCCCGACATCAAGATTCCATCAACCTGATGGTGTAAAGTGC 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 809 TGTCTTTAGATGCTGACTCCGGGCCCAAGGTCAACGGCAATCGCGCCCTGACGCTGC 868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 235 TGATGCTAGAAAGATGAAATATGGGGCCAGAGAGTGTCCGACATGGGGGCTTACACTGC 294
Qy 869 GCAACTTCTGTGCTGGCAGAAAGCTGACAAAGTGAATGACAGACACCCCGAGTACT 928
Db 295 GTAACCTTCTCAACTGGCAGCGCGTTTCAACAGAGCCCAAGACCCCGAGAGCACT 354
Qy 929 GGGACACTGCATCTCTTCCACAGGAGGAGCTGTGTG3---AGCCACCACTGTGTGACA 985
Db 355 ACNACAGGCGCATCTCTCTACACAGACAGACTTCTGTGGCAGGAGGGCTGTGTGACA 414
Qy 986 CCCTGGGCAATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGCTGTGTGATTG 1045
Db 415 CCCTGGGTGTGGCAGACATCGGAGCACTTGTGACCCCAACAAAGCTGCTCGGTGATCG 474
Qy 1046 AGGAGATGGGCTTCCATCAGCTTCCACACTGCCACAGAGCTGGCCAGCTGTTCACAA 1105
Db 475 AGGATAGGGGCTCCAGGCGGCCACACCTGTGGCCCATGAACATAGGCGACGTCCTCAGCA 534
Qy 1106 TGCCCCATGACAATGTGAAAGTCTGTGAGGAGGTGTTTGGAGACTCCGAGGCCAACCA 1165
Db 535 TGCCCCAGCAGACTCCCAAGCCCTGCACACGGCTCTTCGSGGCCATGGCAAGCACACG 594
Qy 1166 TGATGTCGCCGACCTCATCCAGATGACCGTGCACACCCCTGGTCAAGCTGCAGTGTG 1225
Db 595 TGATGGACCGCTGTGTGTCACCTGAACACAGAGAGCTGGCCCTGCTCCCTTGACGGCCA 654
Qy 1226 CCATCATCAGCGACTTCTCTGGACAGCGG 1254
Db 655 TGTTCTCAGGCTGCCACTGCAGGGTGG 683

RESULT 13

US-09-392-184-5/c

; Sequence 5, Application US/09392184

; Patent No. 6395889

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN

; FILE REFERENCE: 5800-55

; CURRENT APPLICATION NUMBER: US/09/392,184

; CURRENT FILING DATE: 1999-09-09

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 5357

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(5357)

; OTHER INFORMATION: repolysin (ADAM family of metalloprotease)

; NAME/KEY: misc_feature

; LOCATION: (1)...(5357)

; OTHER INFORMATION: n = A,T,C or G

US-09-392-184-5

Query Match

Best Local Similarity 8.6%; Score 245.6; DB 4; Length 5357;

Matches 926; Conservative 0; Mismatches 779; Indels 83; Gaps 13;

Qy 630 CAAGCGTTTGTGTCATCCCGGAGTGGAGACCTGGTGGTCCGCGACAGTCAAT 689
Db 4617 CCAGCGTTCGTCAGCAAGAGAGTGGGTGGAGACCTGGTGTAGTATGCCAAAAT 4558
Qy 690 GGTCAAGTTCACAGG---CGCGGACCTGGAAACATTAATCTGCTACGCTGTGGCAACGGC 746
Db 4557 GGTGAGTACACAGCAGCGGAGGTGAGAGCTATGTGCTGACCATCATGACATGGT 4498
Qy 747 GCGCGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCGTTGTGTCAAGGT 806
Db 4497 GGCTGGCTGTTCATGACCCAGCATTTGGGAACCCCATCCACATCACCATTGTGGCCT 4438


```
Db 3404 GTACTTTGCGAGAAAGCTGCGGGACGCGCTGGTGCATGGCACCCCTGCTACAGGTCCG 3345
QY 1935 CACCTCCG-----TCTGTGTCCAAAGCAAGTGCATCAAGCGCTGGCTGTGTAGTGGAACTT 1988
Db 3344 AGCCAGCGGGACCTTCTGCATCAACGGCATCTGTAAGAAGCACTGGGCTGTGACTTTCGAGAT 3285
QY 1989 GGGCTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAAATAGAGCTGCAA 2048
Db 3284 TGACTCCGGTGTATGGAGACCGCTGTGGTGTGTGCCACGGCAACCGCTCCACCTGCCA 3225
QY 2049 GAAGGTGACTGGACTTTCACAAAGCCCATGCATGGCTACAAATTCGTGGTGGCCATCCC 2108
Db 3224 CACCGTAGCGGGACCTTCGAGGAGCGCGAGGGCTGGGTATGTGGATGTGGGGCTGAT 3165
QY 2109 CGCAGCGCCCTCAAGCATCGACATCCGCCAGCGGGTTCACAAAGGGTGTATCGGGGATGA 2168
Db 3164 CCCAGCGCGCGCA-----CGGAGATCCGCATCCAAAGAGGTTGCGGAGGTGC 3117
QY 2169 CAACTTACTTGGCTCTGAAGAAGCAAGCAAGGCAAGTACCTGCTCAACGGGTATTTCTGTGT 2228
Db 3116 CAACTTCTCGCACTGGGAGTGGAGCCCGGAGGAGTACTTCTCAATGTGGCTGGAC 3057
QY 2229 GTCGGCGGTGAGCGGGACCTTGTGTGTAAGGGCAGTCTGTGCGGTACAGCGGCACGGG 2288
Db 3056 CATCCAGTGAACGGGACTACCAAGTGGCAGGGACCACTTCACATACGACGCGAGGG 2997
QY 2289 CACAGCGGTGAGAGCTGCAGGCTTCCCGGCCCATCTTGAGCGCGGTGACCGTGGAGGT 2348
Db 2996 CA---ACTGGGAGAACCTCAGTCCCGGGTCCACCAGGAGGCTGTCTCGATCCAGCT 2940
QY 2349 CCTCTCGTGGGGAGATGACACCGCCCGGGTCCGCTACTCTCTTA 2396
Db 2939 GCTGTTCCAGGAGCAACCTGGGGTGCACTAGGATACACCATCCA 2892
```

RESULT 14

```
us-09-369-364A-6
; Sequence 6, Application us/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3218
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(3003)
us-09-369-364A-6
```

```
Query Match 8.4%; Score 238.6; DB 4; Length 3218;
Best Local Similarity 51.3%; Pred. No. 1.3e-44;
Matches 932; Conservative 0; Mismatches 799; Indels 84; Gaps 13;

QY 562 CTGAGCCCTTACAAAGCGGGGGGGGGCTTGGGGGAGAGTCTGATAGCGCGCGAGGTCT 621
Db 646 CTGAGTCTGACGGGAGGCTTGGGAGACGCGCAGTGGCGGGCGCCACGGCTGAGG 705
QY 622 GGGCGCGCAAGGTTTCTGTCTATCCCGCGGTAGCTGAGACGCTGTGTGTGTCGGGAC 681
Db 706 CGTCTACACCGCGGTGCTGACCAAGAGAGTGTGTGAGACCGCTGTGTAGTCTGAT 765
QY 682 GAGTCAATGGTCAAGTTCACGGG---CGCGGACCTGGAACATTTATCTGCTGACGCTGTG 738
```

```
Db 766 GCCAAATGTGTGAGTACCAGGACAGCCGAGTTGAGAGTATGTGCTGACCATCATG 825
QY 739 GCAACGCGCGGGGACTCTACCGCCATCCAGCATCTCAACCCCATCAACALCTGTGTG 798
Db 826 AACATGGTGTGCTGGCTGTTTCATGACCCAGCATTTGGGAACCCCATCCACATCAAT 885
QY 799 GTCAAGGTGCTGCTTCTTAGAGATGTGATCCGGGCCCAGGTACCGGCAATTCGGGGC 858
Db 886 GTGCGGCTGCTGCTGCTGGAAGATGAGGAGGAGACCTAAAGATCAGCACCATGCAJAC 945
QY 859 CTGACGCTGGCAACTTCTGTGCTGCGAGAGAACTGAACAAAGTGTGAGTACACAGJAC 918
Db 946 AACACCTTGAAGAGCTTCTGAAGTGGCAGAAAGCATCAACATGAAGGGGATGTCAT 1005
QY 919 CCGAGTACTGGGACACTGCCATCTCTTACACGAGCAGGACCTGTGTGGAGCCACCA-- 976
Db 1006 CCCTGCACCATGACACTGCCATCTCTTACACGAGAAAGGACCTGTGTGACCCCATGAC 1065
QY 977 ----CCTGTGACACCTTGGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAGC 1032
Db 1066 CGGCCCTGTGAGACCTTGGGACTGTCCCATGTGGCGGCGATGTGCCAGCGCCAGC 1125
QY 1033 TGCTCTGTCATTGAGGACGATGGCTTCCATCAGCCTTCCACCCTGCCCCACGAGTGGC 1092
Db 1126 TGAGCATCAACGAGGACAGCGGCTGCGCTGAGCTTCACTGTAGCCCCACAGCTCGG 1185
QY 1093 CAGCTGTTCAACATGCCCCATGACATGTGAAGTCTGTGAGGAGTGTTCGGAGCTTC 1152
Db 1186 CACAGTTTGGCATTCAGCATGACGGAAGCGGCAATGACTGTGAGCCGTTGGGAACGA 1245
QY 1153 CGAGCCAAACCATGATGTCCCGACCTTCATCCAGATCGACCGTGGCCAAACCTTGTCA 1212
Db 1246 CCTTTTCATC-----ATGTCTCCACAGCTCTGTACGACGCGCTCCCTCACCTGGTCC 1299
QY 1213 GCTGCGAGTGTGCCATCATCACGACTTCCCTGGACAGCGGCGAGCTGACCTGCTCTG 1272
Db 1300 CGCTGCGCGCGCATGATATACAGGTTCTTTCAGCGGTGGGGCTGTGCTGTCGAC 1359
QY 1273 GACCAACCCAGCAAGCCC---ATCTCCCTGCCGAGGATCTCCGGGGCGCCAGCTACACC 1329
Db 1360 GACCTCTCTGCCAAGGACATTTATGACTTCCCTCGGTGCCACCTGGCTCTCTATGAT 1419
QY 1330 CTGAGCCAGCAGTGGAGCTGGCTTTTGGGTGGGCTCCAGCCCTGTCTTACATGAG 1389
Db 1420 GTAAGCCACGAGTGGCGCTTCCAGTACGGGCTTCTCTGCTTCTGCGAGGACATGAT 1479
QY 1390 TA---CTGCACCAAGCTGTGTGTACCGCGGAAGGCAAGGACAGATGTTGCCAGACC 1446
Db 1480 ATGTCTGCCACACTCTGTGTGCTCTGTGGGACCA-----CCTGTCTACCTCC 1527
QY 1447 CGCCACTTCCCTGGGCGCATGGCACAGCTGTGGCGAGGCAAGCTCTGCTCAAGGG 1506
Db 1528 AAGCTGGATGCAAGCTGTGGAGCGCACCGGTGTGGGGAGAAATAGTGTGTCTCAGTGG 1587
QY 1507 GCCTGCTGGAGAGACAAACCTCAACAAGCACAGGTTGGATGTTCTTGGGCCAAATGG 1566
Db 1588 GAGTGGTACCCGTGGGCTTCCGGCCCGAGGCC---GTGGATGTGTGGTGTGTGCTGGCTG 1644
QY 1567 GATCCCTATGCTGCTCGCGCATGTGTGGGGGCTGACAGTGGCCAGAGGCGAG 1626
Db 1645 AGCGCTGGTCCATCTGCTCAGGAGCTGTGGCATGCGGCTACAGAGCGCGGAGCGGAG 1704
QY 1627 TGCACAAACCCACCCCTGCCAAGGGGGCAAGTACTCCGAGGAGTGTGAGGTGAAATAC 1686
Db 1705 TGCAGCAGCTACGCCCAATACAAAGGAGATACTGTGTGGGTGAGCGCAAGCGTTC 1764
QY 1687 CGATCTGCAATCTGGAGCCCTGCCCGAGCTAGCCTCCGGAAGAGCTTCCGGGAGAG 1746
Db 1765 CGCTCTCAACCTGACGGCTGCC-----CTGTGCGCCCGCTCCTCTCCGCCCAATC 1818
QY 1747 CAGTGTGAGGCTTTCACAGGCTACACACAGCAGCAACCGGCTCACTCTCGCCCTTCCCA 1806
Db 1819 CAGTGCAGCCACTTTGAGGCTATGCTCTACAAAGGCGCAGCTGCACA-----CA 1866
```

QY 1807 TGGGTCCCAAGTACTCCGCGTGTCTCCCGGACAAAGTGAAGTCACTCTGCGAGCC 1866
Db 1867 TGGGTCCCGTGTCAATGACGTGAACCCC-----TGCAGCTGCAGTCCCGGCC 1917
QY 1867 AATGGCACTGGGTACTTCTATGTCTGGCAACCAAGGTGGTGGAGCGGACGCTGTCTCT 1926
Db 1918 GCGAATGAGTACTTTGCCAAGAAGCTGCGGACGCTGTGTGATGGCAACCCCTGCTAC 1977
QY 1927 CTTGACTCCACC-----TCCGTCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1978 CAGTCCGAGCCAGCGGACCTCTGCATCAACGGCATCTGAAGAACGTGGGCTGTGAC 2037
QY 1981 GCGAAGCTGGGTCTCAAGAAGAGATTCCGACAAAGTGTGGGTGTGTGGGGAGACAATAAG 2040
Db 2038 TTGAGATTGACTCCCGTGTCTATGGAGACCGCTGTGGTGTGTGCCACGCAACGGCTCC 2097
QY 2041 AGTGTCAAGAAGTGTACTGGACTTTTCAACCAAGCCCATGCATGGGTACAAATTCGTGGTG 2100
Db 2098 ACCTGCCACACCTGAGCGGACCTTCGAGAGGCGCGAGGTCTGGGGTATGTGGATGTG 2157
QY 2101 GCAATCCCGCGAGCGCTCAAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2160
Db 2158 GGCTGATCCCGCGGCGCA-----CGCGAGATCCGATCCCAAGAGGTGTC 2205
QY 2161 GGGGATGACAACCTACCTGGCTCTGAAGAAGCAGCAAGGCAAGTACCTGCTCAACGGGCAT 2220
Db 2206 GAGCTGCCAACTCTCTGGCACTGCGAGGAGGACCCGCGAAGTACTTCTCAATGCT 2265
QY 2221 TTGCTGGTGTGGCGGTGGAGCGGACCTGGTGGTGAAGGGCAGTGTCTGCGGTACAGC 2280
Db 2266 GGCTGGACCATCAGTGAACGGGACTTACAGGTGGCAGGACCACTTTCATACGCA 2325
QY 2281 GGCACGGGACAGCGGTGGAGCGCTGCAGGCTTCCCGCCCATCTCGAGCGGTGACC 2340
Db 2326 CGCAGGGGCA---ACTGGGAGAACCTCACCTCCCGGGTCCCAAGGAGGCTGTCTGG 2382
QY 2341 GTGGAGTCTCTCC 2355
Db 2383 ATCAGGTGCTGCTCC 2397

RESULT 15
US-09-369-364A-3
; Sequence 3, Application us/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1520
; TYPE: DNA
; ORGANISM: homo sapiens ADAMTS-5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (296)
; OTHER INFORMATION: n-T
; NAME/KEY: misc_feature
; LOCATION: (1443)
; OTHER INFORMATION: n-A
; NAME/KEY: misc_feature
; LOCATION: (1457)
; OTHER INFORMATION: n-G
; NAME/KEY: misc_feature
; LOCATION: (1458)
; OTHER INFORMATION: n -T

US-09-369-364A-3

Query Match 7.6%; Score 217.4; DB 4; Length 1520;
Best Local Similarity 56.1%; Pred. No. 6.2e-40;
Matches 457; Conservative 0; Mismatches 346; Indels 12; Gaps 2;
QY 1603 GGCCTGAGCTGGCCAGGAGCAGTGCACCAACCCCAACCCCTGCCAACGGGGCAAGTAC 1662
Db 360 GGAGTGCAGTTTGCTATCGTCACTGTAAACCCCTCTCCAGAAACAACGACGGCTAC 419
QY 1663 TGCAGGAGCAGTGGGTGAATACCGATCTGCAATCTGGAGCGCTCCGCCAGCTCAGCC 1722
Db 420 TGCACAGGAAGAGGGCCATCTACCACTCTCTGAGTCTCATGCCCCCTCC-----CACCC 473
QY 1723 TCCGGAAGAGCTTCCCGGAGGAGCAGTGTGAGGCTTTCAACGGCTACAAACAGCAGCC 1782
Db 474 AATGTAATCATTTCTGTCATGACAGTGTGAGGCCAAAATAGCTATCAGTCTGATGCA 533
QY 1783 AACCGCTCACTCTCGCGGTGGGATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGAC 1842
Db 534 AAGGAGTCAAAACTTTTGTGAATGGTTCCCAATATGCAAGGTGCTCCGCCAGCGAT 593
QY 1843 AAGTCAAGCTCATCTGCCGAGCAATGCGCACTGGCTACTTCTATGTGCTGGCACCCAAAG 1902
Db 594 GTGTGCAAGCTGACTGCAGAGCCCAAGGCACTGGCTACTATGTGGTATTTCTCCAAAG 653
QY 1903 GTGTGGACGCGCAGCTGTGCTCTCTGACTCCACCTCCGCTGTGTGTCAGGCAAGTGC 1962
Db 654 GTGACGATGCACTGATGATGAGCGGTACAGTAATTCGCTCTCGGGGGAAGTGT 713
QY 1963 ATCAAGGCTGCTGTGATGGCACTGGGCTCCCAAGAGAGATTCGACAGTGTGGGGTG 2022
Db 714 GTGAGAACTGGCTGTGACGGCATCATTTGGCTCAAGGTGCAAGTGTGACAGTGTGCGAGTA 773
QY 2023 TGTGGGGGAGACAATAAGAGCTGCAAGAGGTGACTTGGACTCTTTCACCAAGCCCATGCAT 2082
Db 774 TGTGAGAGAGACAACCTCCAGCTGTACAAGATTGTTGGAACCTTTAATAAGAAAGTAAG 833
QY 2083 GGTACAAATTTGTTGGTGGCCATCCCGCAGCGGCTCAAGCATCGACATCCGCCAGGCC 2142
Db 834 GGTACACTGACGTGGTGGAGTTCCTGAAGGGGCAACCCACATAAAGTTTCGACAGTTC 893
QY 2143 GGTACAAAGGCTGATCGGGGATGACAACCTACCTGGCTCTGAAGAACAGCAAGGCAAG 2202
Db 894 AAAGCCAAAGCAGACTAGATTCTCTGCTATTTAGCCCTGAAAAGAAACAGGTGAG 953
QY 2203 TACTGCTCAACGGCATTTGTTGGTGTGCGGTGGAGCGGACCTGGTGGTGAAGGGC 2262
Db 954 TACCTTATCAATGGAAGTACATGATCTCCACTTCAGAGACTATCATTTGACATCAATGGA 1013
QY 2263 AGTGTGCTGGGTACAGCGGACGCGGACAGCGGTGGAGAGCGCTGCA-----GGCTCC 2316
Db 1014 ACAGTCATGAACATATAGCGGTTGGAGCCACAGGGATGACTTCCCTGCAATGGGCTAC 1073
QY 2317 CGGCGCATCTCGAGCGCTGACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGGCC 2376
Db 1074 TCTGCCAAGAGGAATTTCTAATAGTCAGATTTCTGCAACAGAGCCCAACCAATTA 1133
QY 2377 CGGGTCCGCTACTCTCTCTATCTGCCCCAAAGGCC 2411
Db 1134 GATGTCCTGTTATAGCTTTTGTGTTCCCAAGAGTC 1168

Search completed: May 15, 2003, 22:55:08
Job time : 147 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 16:06:39 ; Search time 576 Seconds

(without alignments)
11154.429 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgctctgctggcgcctcct.....gcgtcctgagggcgtgctga 2853

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result	Score	Query Match	Length	DB	ID	Description
1	2853	100.0	2853	24	AAD35569	Human protease cDN
2	2853	100.0	3446	24	AAD35571	Human protease cDN
3	2846.6	99.8	2853	22	AAH41003	Human metalloprote
4	2817.6	98.8	2930	24	ABK12894	Human protease PRT
5	2817.6	98.8	2937	24	AA997182	Human metalloprote
6	1215	42.6	1518	21	AAA95827	Rat metalloprotein
7	1091.2	38.2	1104	24	AAD35570	Human protease cDN
8	1043.6	36.6	1143	21	AAA95831	Human metalloprote
9	959	33.6	966	24	AAD35568	Human protease cDN

10	664.4	23.3	2670	22	AAH20226	Human ADAM-type me
11	664.4	23.3	2670	22	AAH20226	Human ADAM-type me
12	664.4	23.3	3008	22	AAH202166	Human METH2 encodi
13	664.4	23.3	3008	22	AAH202166	Human METH2 encodi
14	634.4	22.2	4180	20	AAC90058	Human METH2 coding
15	634.4	22.2	4180	20	AAH202166	Human METH1 relate
16	628.2	22.0	3261	22	AAC90067	D67076 cDNA clone.
17	628.2	22.0	3261	22	AAH202000	Human METH1 encodi
18	628.2	22.0	4014	19	AAH200057	Human METH1 coding
19	627.2	22.0	2184	20	AAH200057	Human integrin lig
20	627.2	22.0	4712	20	AAH17990	Human ADAMTS-1 cod
21	626.6	22.0	2853	22	AAH20224	Human secreted pro
22	626.6	22.0	4676	20	AAH19955	Human ADAM-type me
23	617.2	21.6	4858	24	ABH86234	Human Tango-71 enc
24	607.2	21.3	2346	21	AAA95821	Human metalloprote
25	598.2	21.0	3126	24	ABH86300	Human metalloprote
26	590.4	20.7	3638	22	AAH20224	Human metalloprote
27	584.8	20.5	4192	20	AAH20224	Murine ADAMTS-8 CD
28	583.2	20.4	4301	22	AAH20224	Human aggrecanase-
29	583.2	20.4	4303	21	AAH20224	Human metalloprote
30	581.6	20.4	4406	22	AAH20224	Human metalloprote
31	581.6	20.4	4407	21	AAH20224	Human metalloprote
32	581.6	20.4	4407	22	AAH20224	Human metalloprote
33	549	19.2	3250	20	AAH20224	Human metalloprote
34	547.4	19.2	3002	22	AAH20224	Murine aggrecan deg
35	547.4	19.2	5530	21	AAH20224	Murine ADAMTS-5 CD
36	528.6	18.5	1878	24	ABH86276	Human aggrecanase co
37	524.6	18.4	2751	21	AAH20224	Rat metalloprote
38	469.8	16.5	2625	22	AAH20224	Murine ADAMTS-9 CD
39	419.2	14.7	5053	24	ABH88916	Human aggrecanase
40	417.6	14.6	5605	24	ABH88915	Human aggrecanase
41	417.6	14.6	5808	24	AAH20224	Human metalloprote
42	417.6	14.6	5949	24	AAH20224	Human metalloprote
43	417.6	14.6	6013	24	AAH20224	Human ADAMTS-SI nu
44	416	14.6	4086	22	AAH20224	Human metalloprote
45	409.6	14.4	5804	22	AAH20224	Human ADAMTS-9 CD

ALIGNMENTS

RESULT 1

AAD35569

ID AAD35569 standard; cDNA; 2853 BP.

XX

AC AAD35569;

XX

DT 26-JUL-2002 (first entry)

XX Human protease cDNA #2.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;

KW high blood pressure; arthritis; connective tissue disorder; infertility;

KW gene therapy; enzyme; gene; ss.

XX Homo sapiens.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WP1; 2002-372123/40.
DR P-PSDB; AAD22541.
XX
PT Novel nucleic acid encoding a human protease, useful as a hybridization
PT probe for screening libraries and assessing gene expression patterns -
XX
PS Claim 1; Page 35-36; 4lpp; English.
XX
CC The present sequence is a cDNA encoding novel human protein (NHP),
CC human protease. NHPs share structural similarity with animal proteases
CC particularly zinc metalloproteases. Sequences of the invention are
CC useful in therapeutic, diagnostic and pharmacogenomic applications.
CC NHP polynucleotides are used as hybridisation probes for screening
CC libraries and assessing gene expression patterns. They can also be
CC used for treating related biological disorders such as obesity, high
CC blood pressure, arthritis, connective tissue disorders and infertility.
CC They are also used in gene therapy.
XX
SQ Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;
Query Match 100.0%; Score 2853; DB 24; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTTCTGCTGGGCAATCAACCTGGCTTCGCGGGCGAACCCTGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGCAATCAACCTGGCTTCGCGGGCGAACCCTGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTTCATCCGACTGGACCCGGACATTAACGSCCGCGCTAC 120
DB 61 CCAGAGCGGAGGTAGTCTTCATCCGACTGGACCCGGACATTAACGSCCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT 180
DB 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT 180
QY 181 CAGGAGACTTTTACCTTACACTGACGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240
DB 181 CAGGAGACTTTTACCTTACACTGACGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240
QY 241 ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTTTCAGACCTGGCAGCG 300
DB 241 ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTTTCAGACCTGGCAGCG 300
QY 301 TGCCTTCTATTCTGGGACGTAAACCGCGAGCGAGCTCTGCTGCTGTGAGCCTGTGC 360
DB 301 TGCCTTCTATTCTGGGACGTAAACCGCGAGCGAGCTCTGCTGCTGTGAGCCTGTGC 360
QY 361 GGGGGCTCCGCGAGCCTTTGGTACCGAGGCGCGAGTATGTATAGCCCGCTGCC 420
DB 361 GGGGGCTCCGCGAGCCTTTGGTACCGAGGCGCGAGTATGTATAGCCCGCTGCC 420
QY 421 AATCTAGCGCGCGCGCGAGCAACAGCCAGCGCACCTTTCAGAGCGCGG 480
DB 421 AATCTAGCGCGCGCGCGAGCAACAGCCAGCGCACCTTTCAGAGCGCGG 480
QY 481 GGTGTTCCGGCGGGCTTCCGAGAGCCCACTCTGCTGCGGGGTGGGCTCGGGCTGG 540
DB 481 GGTGTTCCGGCGGGCTTCCGAGAGCCCACTCTGCTGCGGGGTGGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTACGGGCGCTGGACCCCTTACAGCCGCGGGGGGCGCTTCGGGGAG 600
DB 541 AACCCCGCATCTACGGGCGCTGGACCCCTTACAGCCGCGGGGGGCGCTTCGGGGAG 600
QY 601 AGTCTAGCGCGCGAGGTCTGGGCGGCGCAAGGTTTCTGTCTATCCCGCGGTACGTG 660
DB 601 AGTCTAGCGCGCGAGGTCTGGGCGGCGCAAGGTTTCTGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGCGGCGAGGTCAATGTCAAGTTCACGCGCGGACCTTGGAAACAT 720
DB 661 GAGACGCTGTGTGCGGCGAGGTCAATGTCAAGTTCACGCGCGGACCTTGGAAACAT 720

QY 721 TATCTGTGAGCTGCTGGCAACGGCGGCGGACTTACCGCCATCCAGCATCTCTCAAC 780
DB 721 TATCTGTGAGCTGCTGGCAACGGCGGCGGACTTACCGCCATCCAGCATCTCTCAAC 780
QY 781 CCCATCAACATCTGTTGGTCAAGGTGCTCTTCTTAGAGATCTGACTCCGGGCCCAAG 840
DB 781 CCCATCAACATCTGTTGGTCAAGGTGCTCTTCTTAGAGATCTGACTCCGGGCCCAAG 840
QY 841 GTCACCGCAATCGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAACTGAAC 900
DB 841 GTCACCGCAATCGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAACTGAAC 900
QY 901 AAATGACTGACACAGCACCCGAGTACTGGACACTGCCATCTCTTCCACGACGAGAC 960
DB 901 AAATGACTGACACAGCACCCGAGTACTGGACACTGCCATCTCTTCCACGACGAGAC 960
QY 961 CTGTGTGAGGACCAACCTGTGACACCTGGGATGGCTGATGTGGGTACCATGTGTGAC 1020
DB 961 CTGTGTGAGGACCAACCTGTGACACCTGGGATGGCTGATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAAAGCTGCTGTGCTATTGAGACATGGGCTTCCATCAGCTTCCACACTGCC 1080
DB 1021 CCCAAGAAAGCTGCTGTGCTATTGAGACATGGGCTTCCATCAGCTTCCACACTGCC 1080
QY 1081 CACGAGCTGGCGGACGCTGTTCAACATGCCCATGACATGTGAAAGTCTGTGAGAGGTG 1140
DB 1081 CACGAGCTGGCGGACGCTGTTCAACATGCCCATGACATGTGAAAGTCTGTGAGAGGTG 1140
QY 1141 TTTGGAAAGCTCCGAGCAACACATGATGTCCCGACCTTCATCCAGATCGACCGTGC 1200
DB 1141 TTTGGAAAGCTCCGAGCAACACATGATGTCCCGACCTTCATCCAGATCGACCGTGC 1200
QY 1201 AACCCCTGGTCAAGCTGCTGCTGATCATCAGCTTCTTGGACAGCGGACGGT 1260
DB 1201 AACCCCTGGTCAAGCTGCTGCTGATCATCAGCTTCTTGGACAGCGGACGGT 1260
QY 1261 GACTGCTTCTGGAGCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTCCCGGGGCG 1320
DB 1261 GACTGCTTCTGGAGCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTCCCGGGGCG 1320
QY 1321 AGCTACACCTGAGCAGCAGTGGCTGCTTTTGGCGTGGGCTCCAGCCCTGTGCT 1380
DB 1321 AGCTACACCTGAGCAGCAGTGGCTGCTTTTGGCGTGGGCTCCAGCCCTGTGCT 1380
QY 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGGAGGCAAGGACACAGTGGTGTGC 1440
DB 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGGAGGCAAGGACACAGTGGTGTGC 1440
QY 1441 CAGACCCGCACTTCCCTGGGCGCATGGCAGCTGCTGGCGAGGCAAGCTCTGCGCTC 1500
DB 1441 CAGACCCGCACTTCCCTGGGCGCATGGCAGCTGCTGGCGAGGCAAGCTCTGCGCTC 1500
QY 1501 AAAGGGGCTGCTGGAGAGACAACTCAACAGCAGAGGTTGGATGTTCTTGGGCG 1560
DB 1501 AAAGGGGCTGCTGGAGAGACAACTCAACAGCAGAGGTTGGATGTTCTTGGGCG 1560
QY 1561 AAATGGATTCCTATGCGCCCTGCTGCGACATGTGTGGGGGCGTGCAGCTGCCAGG 1620
DB 1561 AAATGGATTCCTATGCGCCCTGCTGCGACATGTGTGGGGGCGTGCAGCTGCCAGG 1620
QY 1621 AGGCACTGACCAACCCACCTGCCAAGCGGGCAAGTACTGCGAGGGAGTGGGGTG 1680
DB 1621 AGGCACTGACCAACCCACCTGCCAAGCGGGCAAGTACTGCGAGGGAGTGGGGTG 1680
QY 1681 AAATACCGATCTGCAATCTGAGGCGCTTCCAGCGCTACACACAGCCGCTCCTCTGCC 1740
DB 1681 AAATACCGATCTGCAATCTGAGGCGCTTCCAGCGCTACACACAGCCGCTCCTCTGCC 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCACCGCTTACACAGCGCTACACAGCGCTCCTCTGCC 1800
DB 1741 GAGGAGCAGTGTGAGGCTTTCACCGCTTACACAGCGCTTACACAGCGCTCCTCTGCC 1800
QY 1801 GTGGCATGGTCCCAAGTACTTCCGGCGTGTCTCCCGGGAGCAAGTCAAGCTCATCTGC 1860

Db 1801 GTGGCTGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGCAAGTGCAAGCTCATCTGC 1860
QY 1861 CGAGCCAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGGTGGACGGCAGCTG 1920
Db 1861 CGAGCCAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGGTGGACGGCAGCTG 1920
QY 1921 TGCTCTCTGACTTCCACCTCCGCTCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTTCCACCTCCGCTCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGAT 1980
QY 1981 GGGAACTGGCTTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGCTTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGTGAAGTGAAGTGTGCTTCCAAAGCAAGTGCATCAAGCTGGCTGTGAT 2100
Db 2041 AGCTGCAAGAAGTGAAGTGAAGTGTGCTTCCAAAGCAAGTGCATCAAGCTGGCTGTGAT 2100
QY 2101 GGCATCCCGCAGGGCCCTCAAGCATCGACATCCCGCAGCGGTTACAAAGGGCTGATC 2160
Db 2101 GGCATCCCGCAGGGCCCTCAAGCATCGACATCCCGCAGCGGTTACAAAGGGCTGATC 2160
QY 2161 GGGGATGACAATCTGCTGCTGAGAACAGCAAGCAAGTGTGCTTCAACGGGCT 2220
Db 2161 GGGGATGACAATCTGCTGCTGAGAACAGCAAGCAAGTGTGCTTCAACGGGCT 2220
QY 2221 TTCTGGTGTGCGCGGTGAGCGGACCTGTGGTGTGAAGTGTGCTTCAACGGGCT 2280
Db 2221 TTCTGGTGTGCGCGGTGAGCGGACCTGTGGTGTGAAGTGTGCTTCAACGGGCT 2280
QY 2281 GGCACGGGCACAGCGGTGAGAGCTGTGAGGCTTCCCGGCGCATCTGAGCGGCTGACC 2340
Db 2281 GGCACGGGCACAGCGGTGAGAGCTGTGAGGCTTCCCGGCGCATCTGAGCGGCTGACC 2340
QY 2341 GTGGAGTCTCTCGTGGGGAAGATGACACCGCCCGGTTCCGCTATCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCGTGGGGAAGATGACACCGCCCGGTTCCGCTATCTTCTATCTG 2400
QY 2401 CCCAAGAGCTCTGGGAGACAAAGTCTCTATCCCAAGAGCCCGCGGACCTCTGTC 2460
Db 2401 CCCAAGAGCTCTGGGAGACAAAGTCTCTATCCCAAGAGCCCGCGGACCTCTGTC 2460
QY 2461 TTGCACACAGGCTCCTCAGCCTCTCCAAACAGGTGGAGTGGAGCGGACGACGCGCCCT 2520
Db 2461 TTGCACACAGGCTCCTCAGCCTCTCCAAACAGGTGGAGTGGAGCGGACGACGCGCCCT 2520
QY 2521 GCACGCTGGTGGCTGGCAGCTGGGGCGCTGCTCCGGAGCTGGGGCAGTGGCTGAG 2580
Db 2521 GCACGCTGGTGGCTGGCAGCTGGGGCGCTGCTCCGGAGCTGGGGCAGTGGCTGAG 2580
QY 2581 AAGCGGCGGTGGACTGTGCGGGCTCCCGCGGAGCGCAGCGTGGCTGATGCA 2640
Db 2581 AAGCGGCGGTGGACTGTGCGGGCTCCCGCGGAGCGCAGCGTGGCTGATGCA 2640
QY 2641 GCCATCGGCGGTGGAGACACAGCTGTGCGGGAGGCCGCCCGACCTGGAGCTCAGC 2700
Db 2641 GCCATCGGCGGTGGAGACACAGCTGTGCGGGAGGCCGCCCGACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGTCTCAAGAGCTGGCGCGGGATTTAGAGCGCTTACTCAAGTGT 2760
Db 2701 GCCTGGTCAACCTGTCTCAAGAGCTGGCGCGGGATTTAGAGCGCTTACTCAAGTGT 2760
QY 2761 GTGGGCCACGAGCGGCTGCTGGCGGGAGCCAGTGCACACTTGCACCGCAAGCCCGAG 2820
Db 2761 GTGGGCCACGAGCGGCTGCTGGCGGGAGCCAGTGCACACTTGCACCGCAAGCCCGAG 2820
QY 2821 GAGCTGGACTTCTGCTGCTGAGCGCGTGTGTA 2853
Db 2821 GAGCTGGACTTCTGCTGCTGAGCGCGTGTGTA 2853

RESULT 2

AAD35571

ID AAD35571 standard; cDNA; 3446 BP.

XX

AC AAD35571;

XX

DT 26-JUL-2002 (first entry)

XX

DE Human protease cDNA #4.

XX

Human; novel human protein; NHP; protease; biological disorder; obesity; high blood pressure; arthritis; connective tissue disorder; infertility; gene therapy; enzyme; ss.

XX

OS Homo sapiens.

XX

PN WO200226949-A2.

XX

PD 04-APR-2002.

XX

XX 27-SEP-2001; 2001WO-US30350.

XX

XX 29-SEP-2000; 2000US-236689P.

PR

(LEXI-) LEXICON GENETICS INC.

XX

PA Friddle CJ, Hilbun E;

XX

PI WPI; 2002-372123/40.

XX

Novel nucleic acid encoding a human protease, useful as a hybridization probe for screening libraries and assessing gene expression patterns - Disclosure; Page 40-41; 41pp; English.

XX

The present sequence is a cDNA encoding novel human protein (NHP), human protease. NHPs share structural similarity with animal proteases particularly zinc metalloproteases. Sequences of the invention are useful in therapeutic, diagnostic and pharmacogenomic applications. NHP polynucleotides are used as hybridisation probes for screening libraries and assessing gene expression patterns. They can also be used for treating related biological disorders such as obesity, high blood pressure, arthritis, connective tissue disorders and infertility. They are also used in gene therapy.

XX

SQ Sequence 3446 BP; 612 A; 1114 C; 1101 G; 619 T; 0 other;

Query Match 100.0%; Score 2853; DB 24; Length 3446;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 ATGCTTCTGCTGGCATCCCTAACCCCTGGCTTTCGCGGGGGAACCCCTGGAGGCTCTGAG 60

Db

397 ATGCTTCTGCTGGCATCCCTAACCCCTGGCTTTCGCGGGGGAACCCCTGGAGGCTCTGAG 456

QY

61 CCAGAGCGGAGGTAGTGTTCCTCCATCCGACTGGACCGCGGACATTAAACGCCCGCTAC 120

Db

457 CCAGAGCGGAGGTAGTGTTCCTCCATCCGACTGGACCGCGGACATTAAACGCCCGCTAC 516

QY

121 TACTGGGGGTCGCCGAGGACTCCGGGATCAGTCTTTCAGATCAGAGCATTT 180

Db

517 TACTGGGGGTCGCCGAGGACTCCGGGATCAGTCTTTCAGATCAGAGCATTT 576

QY

181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTTCAGATCAGAGCATTT 240

Db

577 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTTCAGATCAGAGCATTT 636

QY

241 ACTGAGCATCTGGCGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACCTGGGACGC 300

Db

637 ACTGAGCATCTGGCGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACCTGGGACGC 696

QY

301 TGCTTCTATTTGGGGACGTGAACCGCGGACTCGTTCGCTGCTGTGAGCGCTGTC 360

Db

697 TGCTTCTATTTGGGGACGTGAACCGCGGACTCGTTCGCTGCTGTGAGCGCTGTC 756

QY 361 GGGGGCTCGGGAGCCTTTGGCTACCGAGCGCGAGTATGTCTATAGCCCGCTGCC 420
DB 757 GGGGGCTCGGGAGCCTTTGGCTACCGAGCGCGAGTATGTCTATAGCCCGCTGCC 816
QY 421 AATGCTAGCGCGCGCGGCGGAGGCAACAGCAAGCGGCGGCAACCTTCTCCAGCGCGG 480
DB 817 AATGCTAGCGCGCGCGGCGGAGGCAACAGCAAGCGGCGGCAACCTTCTCCAGCGCGG 876
QY 481 GGTGTTCCGGGGGGGCTTTCCGGAGACCCACCTCTCGCTGCGGGTGCCCTCGGGCTGG 540
DB 877 GGTGTTCCGGGGGGGCTTTCCGGAGACCCACCTCTCGCTGCGGGTGCCCTCGGGCTGG 936
QY 541 AACCCGCCATCTACGGGCGCTTGAACCTTACAAGCGGCGGCGGCGGCTTTCCGGGAG 600
DB 937 AACCCGCCATCTACGGGCGCTTGAACCTTACAAGCGGCGGCGGCGGCTTTCCGGGAG 996
QY 601 AGTCGTAGCGCGCGGAGGCTTGGGCGCGGCAAGCGCTTCTATCCCGGGTACGTCG 660
DB 997 AGTCGTAGCGCGCGGAGGCTTGGGCGCGGCAAGCGCTTCTATCCCGGGTACGTCG 1056
QY 661 GAGACGCTGCTGTCGCGGACGAGTCAATGGTCAAGTTCACGGCGGCGGACCTGGAAAT 720
DB 1057 GAGACGCTGCTGTCGCGGACGAGTCAATGGTCAAGTTCACGGCGGCGGACCTGGAAAT 1116
QY 721 TATCTGTGACGCTGCTGCAACGGCGGCGGCACTCTACCGGCATCCCAAGCATCTCAAC 780
DB 1117 TATCTGTGACGCTGCTGCAACGGCGGCGGCACTCTACCGGCATCCCAAGCATCTCAAC 1176
QY 781 CCCATCAACATCTGTTGGTCAAGGTGCTGCTTTAGAGATCGTACTCCGGCGGCGCAAG 840
DB 1177 CCCATCAACATCTGTTGGTCAAGGTGCTGCTTTAGAGATCGTACTCCGGCGGCGCAAG 1236
QY 841 GTACCGGCAATGGCGCTGACGCTGCGCAACTTCTGTGCGTGGCAGAAAGCTGAAC 900
DB 1237 GTACCGGCAATGGCGCTGACGCTGCGCAACTTCTGTGCGTGGCAGAAAGCTGAAC 1296
QY 901 AAAGTGAAGTGAACGACCCCGAGTACTGGGACACTGCCATCCTCTTACACGAGGAC 960
DB 1297 AAAGTGAAGTGAACGACCCCGAGTACTGGGACACTGCCATCCTCTTACACGAGGAC 1356
QY 961 CTGTGTGGACCCACCATCTGTACACCTGGGATGGCTATGTGGGTACCATGTGTGAC 1020
DB 1357 CTGTGTGGACCCACCATCTGTACACCTGGGATGGCTATGTGGGTACCATGTGTGAC 1416
QY 1021 CCCAAGAAAGCTGCTGTCTATTGAGGAGATGGGCTTCCATCAGCCTTCAACACTGCC 1080
DB 1417 CCCAAGAAAGCTGCTGTCTATTGAGGAGATGGGCTTCCATCAGCCTTCAACACTGCC 1476
QY 1081 CACGAGCTGGGCGACCGTGTTCACATGCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
DB 1477 CACGAGCTGGGCGACCGTGTTCACATGCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1536
QY 1141 TTTGGGAAGCTCCGAGCCACCATCATGTCCCGACCTCATCCAGATCGACCGTGGC 1200
DB 1537 TTTGGGAAGCTCCGAGCCACCATCATGTCCCGACCTCATCCAGATCGACCGTGGC 1596
QY 1201 AACCCCTGGTCAGCTGCTGATGATGATCATCATCAGGACTTCTTGGACAGCGGCGACG 1260
DB 1597 AACCCCTGGTCAGCTGCTGATGATGATCATCATCAGGACTTCTTGGACAGCGGCGACG 1656
QY 1261 GACTGCTCTGTGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGCCGGCGGC 1320
DB 1657 GACTGCTCTGTGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGCCGGCGGC 1716
QY 1371 AGCTACACCTGTAGCCAGAGTGGAGCTTGTGGCTGGGCTTCCAGGCGCTGTCT 1380
DB 1717 AGCTACACCTGTAGCCAGAGTGGAGCTTGTGGCTGGGCTTCCAGGCGCTGTCT 1776
QY 1381 TACATGAGTACTGACCAAGCTGTGTGACCCGGGAAGGCCAAGGAGACAGATGGTGTGC 1440
DB 1777 TACATGAGTACTGACCAAGCTGTGTGACCCGGGAAGGCCAAGGAGACAGATGGTGTGC 1836

QY 1441 CAGACCGCGCATTCCCTGGGCGGATGGACACAGCTGTGGCGAGGCAAGCTCTGCCIC 1500
DB 1837 CAGACCGCGCATTCCCTGGGCGGATGGACACAGCTGTGGCGAGGCAAGCTCTGCCIC 1896
QY 1501 AAAGGGCGCTCGGTGGAGAGACACAACCTCAACAAGCACAGGTGGATGTTCTCTGGGC 1560
DB 1897 AAAGGGCGCTCGGTGGAGAGACACAACCTCAACAAGCACAGGTGGATGTTCTCTGGGC 1956
QY 1561 AAATGGGATCCCTATGTCCTGCTCGCGACATGTGGTGGGGCGGTGCACTGGCAAG 1620
DB 1957 AAATGGGATCCCTATGTCCTGCTCGCGACATGTGGTGGGGCGGTGCACTGGCAAG 2016
QY 1621 AGCAGTGCACCAACCCACCTGTCACAGGGGCGAAGTACTGCGAGGAGTGAAGGIG 1680
DB 2017 AGCAGTGCACCAACCCACCTGTCACAGGGGCGAAGTACTGCGAGGAGTGAAGGIG 2076
QY 1681 AAATACCGATCTGCAATCTGGAGCCTGTCACAGCTGCACTCCCGAAGAGCTTCGG 1740
DB 2077 AAATACCGATCTGCAATCTGGAGCCTGTCACAGCTGCACTCCCGAAGAGCTTCGG 2136
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACAGCAGCACCAACCGGCTCACTTCGCC 1800
DB 2137 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACAGCAGCACCAACCGGCTCACTTCGCC 2196
QY 1801 GTGGATGGGTGCCCAAGTACTCGGGGTGCTCCCCGGGACAGTGCACAGCTCATCTGC 1860
DB 2197 GTGGATGGGTGCCCAAGTACTCGGGGTGCTCCCCGGGACAGTGCACAGCTCATCTGC 2256
QY 1861 CGAGCAATGCACTGGCTACTTCTATGTGTGCAACCCCAAGTGGTGGAGCGGACGCTG 1920
DB 2257 CGAGCAATGCACTGGCTACTTCTATGTGTGCAACCCCAAGTGGTGGAGCGGACGCTG 2316
QY 1921 TGCTCTCTGACTCCACTCCGCTGCTGTCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
DB 2317 TGCTCTCTGACTCCACTCCGCTGCTGTCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 2376
QY 1981 GGGACCTGGCTCCAGAGAGATTCGCAAGTGTGGGTGCTGGGGAGAGCAATAAG 2040
DB 2377 GGGACCTGGCTCCAGAGAGATTCGCAAGTGTGGGTGCTGGGGAGAGCAATAAG 2436
QY 2041 AGCTCAAGAAAGTGAAGTGTGCTTCAAGGCAAGTGCATCAAGTTCGTTGGTG 2100
DB 2437 AGCTCAAGAAAGTGAAGTGTGCTTCAAGGCAAGTGCATCAAGTTCGTTGGTG 2496
QY 2101 GCATPCCCGCAGGCGCTCAAGCATCGCATCCGAGCGCGGTTTCAAGAGGCTCATC 2160
DB 2497 GCATPCCCGCAGGCGCTCAAGCATCGCATCCGAGCGCGGTTTCAAGAGGCTCATC 2556
QY 2161 GGGGATGACAATCTCTGGCTGTGAAGAAACAGCAAGCAAGTACCTGCTCAACGGGAT 2220
DB 2557 GGGGATGACAATCTCTGGCTGTGAAGAAACAGCAAGCAAGTACCTGCTCAACGGGAT 2616
QY 2221 TTTGCTGTGTGGGGTGGAGCGGACCTGGTGTGAAGGCGAGTGTGCTGCGGTACAGC 2280
DB 2617 TTTGCTGTGTGGGGTGGAGCGGACCTGGTGTGAAGGCGAGTGTGCTGCGGTACAGC 2676
QY 2281 GGCACGGGACAGCGGTGGAGCGCTGAGGCTTCCCGGCGCCATCCTGGAGCGGCTGACC 2340
DB 2677 GGCACGGGACAGCGGTGGAGCGCTGAGGCTTCCCGGCGCCATCCTGGAGCGGCTGACC 2736
QY 2341 GTGGAGGTCTCTCGGTGGGAAGATGACACCGCGCGGGTCCGCTACTCTCTATCTG 2400
DB 2737 GTGGAGGTCTCTCGGTGGGAAGATGACACCGCGCGGGTCCGCTACTCTCTATCTG 2796
QY 2401 CCCAAGAGCCTCGGAGAGCAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTC 2460
DB 2797 CCCAAGAGCCTCGGAGAGCAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTC 2856
QY 2461 TTTGCAACAGCGCTCTCAGCCTCTCCAGAGTGGAGCGAGCGAGAGCGGCGCCCT 2520
DB 2857 TTTGCAACAGCGCTCTCAGCCTCTCCAGAGTGGAGCGAGCGGAGAGCGGCGCCCT 2916
QY 2521 GCACGCTGGGTGGCTGGAGCTGGGGGCGGCTGCTCCGCGAGCTGCGGAGTGGGCTGCAG 2580

Db 2917 GCACGCTGGTGGCTGGCAGCTGGGGCGGTGCTCCGCGAGCTGGGCGAGTGGCTGCAG 2976
QY 2581 AAGCGGGGGTGGAGTGTCTGGGGCTCCCGCGGAGCGACGGTCCCTGTGTATCCA 2640
Db 2977 AAGCGGGGGTGGAGTGTCTGGGGCTCCCGCGGAGCGACGGTCCCTGTGTATCCA 3036
QY 2641 GCCCATCGCCGCTGGAGACACAAGCCCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
Db 3037 GCCCATCGCCGCTGGAGACACAAGCCCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 3096
QY 2701 GCCTGGTCAACCTGTCTCAAGAGTGTGCGCCGGGATTTACAGCGGCTCACTCAAGTGT 2760
Db 3097 GCCTGGTCAACCTGTCTCAAGAGTGTGCGCCGGGATTTACAGCGGCTCACTCAAGTGT 3156
QY 2761 GTGGCCACGAGCGGCTGTCTGGCCGGGACCAAGTGTCACTTGCACCGCCAG 2820
Db 3157 GTGGCCACGAGCGGCTGTCTGGCCGGGACCAAGTGTCACTTGCACCGCCAG 3216
QY 2821 GAGCTGGACTTCTGCTCTGAGCGCGTGTGA 2853
Db 3217 GAGCTGGACTTCTGCTCTGAGCGCGTGTGA 3249

RESULT 3

AAH41003

ID AAH41003 standard; cDNA; 2853 BP.

XX AC AAH41003;

XX DT

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

CC osteopathic and antiarthritic activity. The present sequence represents
cDNA encoding the metalloprotease termed MDT56.

XX
SQ Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;

Query Match 99.8%; Score 2846.6; DB 22; Length 2853;
Best local Similarity 99.9%; Pred. No. 0;
Matches 2849; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCCGGGGAACCGCTGGAGCTCTGAG 60
Db 1 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCCGGGGAACCGCTGGAGCTCTGAG 60

QY 61 CCAGAGCGGAGTAGTTCGTTCCATCCGACTGGACCCGACATTAACGCGCGCCGATAC 120
Db 61 CCAGAGCGGAGTAGTTCGTTCCATCCGACTGGACCCGACATTAACGCGCGCCGATAC 120

QY 121 TACTGGGGGTCGCCGAGGACTCCGGGATCAGGACTATTTTTCAGATCACAGCATTT 180
Db 121 TACTGGGGGTCGCCGAGGACTCCGGGATCAGGACTATTTTTCAGATCACAGCATTT 180

QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCGCTTC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCGCTTC 240

QY 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTCACCGGGGCTCTTCAGACCTCGCACGC 300
Db 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTCACCGGGGCTCTTCAGACCTCGCACGC 300

QY 301 TGCTTCTATTCGGGACGTGAACGCCGAGCGGACTCCTTCGCTGCTGAGCCTGTGC 360
Db 301 TGCTTCTATTCGGGACGTGAACGCCGAGCGGACTCCTTCGCTGCTGAGCCTGTGC 360

QY 361 GGGGGCTCCGGGAGCTTTGGCTACCGAGCGCGGAGTATGCTATTAGCCGCTGCC 420
Db 361 GGGGGCTCCGGGAGCTTTGGCTACCGAGCGCGGAGTATGCTATTAGCCGCTGCC 420

QY 421 AATGCTAGCGCGCGCGCGCGCACACAGCGGCGCACACCTTCTCCAGCGCGCG 480
Db 421 AATGCTAGCGCGCGCGCGCGCACACAGCGGCGCACACCTTCTCCAGCGCGCG 480

QY 481 GGTGTTCCGGGGGGCTTCGGGAGACCCACCTCTCGTGGGGGTGCGCTGCGGCTGG 540
Db 481 GGTGTTCCGGGGGGCTTCGGGAGACCCACCTCTCGTGGGGGTGCGCTGCGGCTGG 540

QY 541 AACCCCGCATCTACGGGCTTGGACCTTACAGCGCGCGGCGGCTTCGGGAG 600
Db 541 AACCCCGCATCTACGGGCTTGGACCTTACAGCGCGCGGCGGCTTCGGGAG 600

QY 601 AGTCTAGCGCGCGCGCGCGCGCACAGCGGCTTCTGCTGCTATCCCGGCTACG 660
Db 601 AGTCTAGCGCGCGCGCGCGCGCACAGCGGCTTCTGCTGCTATCCCGGCTACG 660

QY 661 GAGACGCTGGTGTGCGGAGCAGTCAATGTCCACGGCGCGGACCTGGAACAT 720
Db 661 GAGACGCTGGTGTGCGGAGCAGTCAATGTCCACGGCGCGGACCTGGAACAT 720

QY 721 TATCTGCTGACGCTGTGCAACGCGGCGGAGTCTACCGCCATCCAGCATCTCAAC 780
Db 721 TATCTGCTGACGCTGTGCAACGCGGCGGAGTCTACCGCCATCCAGCATCTCAAC 780

QY 781 CCATCAACATCGTTGTGCTCAAGGTGCTTCTTAGAGATCGTACTCCGGGCGCCAG 840
Db 781 CCATCAACATCGTTGTGCTCAAGGTGCTTCTTAGAGATCGTACTCCGGGCGCCAG 840

QY 841 GTCACCGGAGTGGGCGCTGACGCTGGCGCACTTCTGTGCTGGCAGAGAAGCTGAAC 900
Db 841 GTCACCGGAGTGGGCGCTGACGCTGGCGCACTTCTGTGCTGGCAGAGAAGCTGAAC 900

QY 901 AAAGTGTGAGACACCCCGGAGTCTGGGAGCTGCTTCTTCCAGGAGGAGAC 960
Db 901 AAAGTGTGAGACACCCCGGAGTCTGGGAGCTGCTTCTTCCAGGAGGAGAC 960


```
QY 961 CTGTGTGGAGCCACCACCTGTGTGACACCCCTGGGCATGGCTGTATGTGGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCACCTGTGTGACACCCCTGGGCATGGCTGTATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGCTGCTGTCTGTCATTTGAGGAGGATGGGCTTCCATCAGCCCTTACCACCTGCC 1080
Db 1021 CCCAAGAGAAGCTGCTGTCTGTCATTTGAGGAGGATGGGCTTCCATCAGCCCTTACCACCTGCC 1080
QY 1081 CAGAGCTGGGCGACACCTGTTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CAGAGCTGGGCGACACCTGTTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACCAACACATGATGTGTCCTCCGACCCCTCATPCAGATPCAGCCGTGCC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACCAACACATGATGTGTCCTCCGACCCCTCATPCAGATPCAGCCGTGCC 1200
QY 1201 AACCCCTGGTCAAGCTGCTGAGTGTGCTGATCATCACCAGCTTCTTGACACAGGGGACGGT 1260
Db 1201 AACCCCTGGTCAAGCTGCTGAGTGTGCTGATCATCACCAGCTTCTTGACACAGGGGACGGT 1260
QY 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCCCGAGGATCTGCCGGGGGCC 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCCCGAGGATCTGCCGGGGGCC 1320
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTTCCAGCCCTGTCT 1380
Db 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTTCCAGCCCTGTCT 1380
QY 1381 TACATGCACTGTGACCAAGCTGTGGTGACCGGGAAGGCTTCAAGGACAGATGTGTG 1440
Db 1381 TACATGCACTGTGACCAAGCTGTGGTGACCGGGAAGGCTTCAAGGACAGATGTGTG 1440
QY 1441 CAGACCCGCCATCTCCCTGGGCGGATGTGCACCAAGTGTGGCGAGGGCAAGCTTCTGCCCTC 1500
Db 1441 CAGACCCGCCATCTCCCTGGGCGGATGTGCACCAAGTGTGGCGAGGGCAAGCTTCTGCCCTC 1500
QY 1501 AAAGGGCCCTGGTGAGAGACACACCTCAACAAGCAGAGTGTGATGGTCTTCTGGGCC 1560
Db 1501 AAAGGGCCCTGGTGAGAGACACACCTCAACAAGCAGAGTGTGATGGTCTTCTGGGCC 1560
QY 1561 AAATGGGATCCCTATGAGCCCTGTGTCGACACATGTGTGGGGGCGGTGCAAGCTGGCCAGG 1620
Db 1561 AAATGGGATCCCTATGAGCCCTGTGTCGACACATGTGTGGGGGCGGTGCAAGCTGGCCAGG 1620
QY 1621 AGGCAGTGCACCAACCCCTGCTGCGCACATGTGTGGGGGCGGTGCAAGCTGGCCAGG 1680
Db 1621 AGGCAGTGCACCAACCCCTGCTGCGCACATGTGTGGGGGCGGTGCAAGCTGGCCAGG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCCCTCCGAAAGAGCTTCCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCCCTCCGAAAGAGCTTCCGG 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTACACCAACAGCACCACCGGCTTCACTCTCGCC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTACACCAACAGCACCACCGGCTTCACTCTCGCC 1800
QY 1801 GTGGCATGGTGCCCAAGTACTCGGGGTGTCTCCCGGGACAAAGTGCAGCTCACTCTGC 1860
Db 1801 GTGGCATGGTGCCCAAGTACTCGGGGTGTCTCCCGGGACAAAGTGCAGCTCACTCTGC 1860
QY 1861 CGAGCCATGGCACTGGCTACTTCTATGTGTGTCGACCCCAAGGTGTGTGACGGCAGGCTG 1920
Db 1861 CGAGCCATGGCACTGGCTACTTCTATGTGTGTCGACCCCAAGGTGTGTGACGGCAGGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCGCTGTGTGTCACAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCGCTGTGTGTCACAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGGACCTGGCTCCAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGACCTGGCTCCAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGCTAGCTACAAATTCGTGGTG 2100
```

```
Db 2041 AGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGCTAGCTACAAATTCGTGGTG 2100
QY 2101 GCCATCCCGCAGCGCCTCAAGCATCCGCCAGCGCGGTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGCAGCGCCTCAAGCATCCGCCAGCGCGGTACAAAGGCTGATC 2160
QY 2161 GGGGATGACAACCTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGCAT 2220
Db 2161 GGGGATGACAACCTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGCAT 2220
QY 2221 TTGCTGGTGTGCGCGGTGAGCGGACCTGTGTGAGGGCAGTCTGCTGCGGTACAGC 2280
Db 2221 TTGCTGGTGTGCGCGGTGAGCGGACCTGTGTGAGGGCAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCAGGGCAGCAGCGGTGGAGAGCTGCAGGCTTCCCGGGCCATCCTGGAGCCCTGACC 2340
Db 2281 GGCAGGGCAGCAGCGGTGGAGAGCTGCAGGCTTCCCGGGCCATCCTGGAGCCCTGACC 2340
QY 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGSAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTC 2460
Db 2401 CCCAAGAGCCTCGSAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTC 2460
QY 2461 TTGCAACAACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGAGCCGCCCT 2520
Db 2461 TTGCAACAACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGAGCCGCCCT 2520
QY 2521 GCACCTGGGTGGTGGCAGCTGGGGGCGGTCTCCCGGAGCTGCGGAGTGGCCTGCG 2580
Db 2521 GCACCTGGGTGGTGGCAGCTGGGGGCGGTCTCCCGGAGCTGCGGAGTGGCCTGCG 2580
QY 2581 AAGCGGGCGGTGGAGTGTGCGGGGTCTCCCGGGGAGCGCAGCGCTGCTGTATGCA 2640
Db 2581 AAGCGGGCGGTGGAGTGTGCGGGGTCTCCCGGGGAGCGCAGCGCTGCTGTATGCA 2640
QY 2641 GCCCATCGGCGGTGGAGACACAGCCCTGCGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
Db 2641 GCCCATCGGCGGTGGAGACACAGCCCTGCGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
QY 2701 GCTGTGTACCTGTCTCAAGAGCTGCGGCGGGGATTTTACAGGCGCTCACTCAAGTGT 2760
Db 2701 GCTGTGTACCTGTCTCAAGAGCTGCGGCGGGGATTTTACAGGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCCAGGAGCGGCTGTGTGGCCCGGGACCAAGTGCAGCTTGCACCGCAAGCCGAG 2820
Db 2761 GTGGGCCAGGAGCGGCTGTGTGGCCCGGGACCAAGTGCAGCTTGCACCGCAAGCCGAG 2820
QY 2821 GAGCTGGACTTCTGCGTCTGAGGCGGTGCTGA 2853
Db 2821 GAGCTGGACTTCTGCGTCTGAGGCGGTGCTGA 2853
```

RESULT 4
ABK12894

ID ABK12894 standard; cDNA; 2930 BP.

XX ABK12894;

XX 09-APR-2002 (first entry)

XX Human protease PRTS-11 cDNA sequence.

XX Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
XX cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
XX inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
XX cell proliferative disorder; developmental disorder; epilepsy;
XX Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
XX reproductive disorder; endometriosis; ss.

QY 1321 AGCTACACCTGAGCCAGCAGTCCGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCT 1380
DB 1395 AGCTACACCTGAGCCAGCAGTCCGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCT 1454
QY 1381 TACATGAGTACTGACCAAGCTGTGGTCCACCGGAAGGCAAGGACATGTGTGTG 1440
DB 1455 TACATGAGTACTGACCAAGCTGTGGTCCACCGGAAGGCAAGGACATGTGTGTG 1514
QY 1441 CAGACCGCCACATTCCTCGGCGGATGAGACAGCTGTGGGAGGGAAGTCTGTGCTC 1500
DB 1515 CAGACCGCCACATTCCTCGGCGGATGAGACAGCTGTGGGAGGGAAGTCTGTGCTC 1574
QY 1501 AAGGGGCTCGCTGGAGAGACAACTTCAACAACAGAGGTGTGTCTTCTGGGCG 1560
DB 1575 AAGGGGCTCGCTGGAGAGACAACTTCAACAACAGAGGTGTGTCTTCTGGGCG 1634
QY 1561 AAATGGGATCCCTATGSCCCCTCTCGCGACATGTGTGGGGCGTGCAGCTGGCCAG 1620
DB 1635 AAATGGGATCCCTATGSCCCCTCTCGCGACATGTGTGGGGCGTGCAGCTGGCCAG 1694
QY 1621 AGGAGTGCACCAACCCACCCCTGCCAAGGGGGCAAGTACTGCGAGGAGTGGGGTG 1680
DB 1695 AGGAGTGCACCAACCCACCCCTGCCAAGGGGGCAAGTACTGCGAGGAGTGGGGTG 1754
QY 1681 AAATACCGATCCCTGCAATCTGGAGCCCTGCCCGACATGTGTGGGGCGTGCAGCTTCCGG 1740
DB 1755 AAATACCGATCCCTGCAATCTGGAGCCCTGCCCGACATGTGTGGGGCGTGCAGCTTCCGG 1814
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGGTACACAGCAGCAACACCGGCTCACTTCGGC 1800
DB 1815 GAGGAGCAGTGTGAGGCTTTCAACGGGTACACAGCAGCAACACCGGCTCACTTCGGC 1874
QY 1801 GTGGCATGGTGGCCCAAGTACTCGGGTGTCTCCCGGACAAAGTGCAGCTCACTTCG 1860
DB 1875 GTGGCATGGTGGCCCAAGTACTCGGGTGTCTCCCGGACAAAGTGCAGCTCACTTCG 1934
QY 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTCTGTGCTGGCAACCCAA---GGTGGTGGACGCAAG 1917
DB 1935 CGAGCCAAATGGCACTGGCTACTTCTATGTCTGTGCTGGCAACCCAAAGTGGTGGACGCAAG 1994
QY 1918 CTGTGCTCTCTGACTCCACTCGCTGTGTCTCAAGGCAAGTGCAGCTCAAGCTGGCTGT 1977
DB 1995 CTGTGCTCTCTGACTCCACTCGCTGTGTCTCAAGGCAAGTGCAGCTCAAGCTGGCTGT 2054
QY 1978 GATGGAACTGGGCTCCAAAGAGATTCGAAAGTGGGGTGTGTGGGGGAGACAAT 2037
DB 2055 GATGGAACTGGGCTCCAAAGAGATTCGAAAGTGGGGTGTGTGGGGGAGACAAT 2114
QY 2038 AAGAGCTGCAAGAGGTGACTGGACTCTTCAACCAAGCCCATGATGCTACAATTTCTGT 2097
DB 2115 AAGAGCTGCAAGAGGTGACTGGACTCTTCAACCAAGCCCATGATGCTACAATTTCTGT 2174
QY 2098 GTGGCCATCCCGCAGCGGCTCAAGCATGCATCCGCTAGCGCGGTACAAAGGCTG 2157
DB 2175 GTGGCCATCCCGCAGCGGCTCAAGCATGCATCCGCGCGGCTTACAAAGGCTG 2234
QY 2158 ATCGGGATGCAACTACCTACCTGGCTCTGAAGACAGCAAGCAAGTACCTGCTCAAGCGG 2217
DB 2235 ATCGGGATGCAACTACCTACCTGGCTCTGAAGACAGCAAGCAAGTACCTGCTCAAGCGG 2294
QY 2218 CATTTCTGTGTGCGGGTGGAGCGGACCTGGTGTGTAAGGCGAGTCTGTGCGGTAC 2277
DB 2295 CATTTCTGTGTGCGGGTGGAGCGGACCTGGTGTGTAAGGCGAGTCTGTGCGGTAC 2354
QY 2278 AGCGCACGGGCACAGCGGTGGAGAGCTTCAGAGCTTCCCGGCCATCTTGGAGCGCTG 2337
DB 2355 AGCGCACGGGCACAGCGGTGGAGAGCTTCAGAGCTTCCCGGCCATCTTGGAGCGCTG 2414
QY 2338 ACCGTGAGGTCTCTCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTAT 2397
DB 2415 ACCGTGAGGTCTCTCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTAT 2474

QY 2398 CTGCCAAAGAGCCTCGGAGGACAAAGTCTCTATCCCAAGGACCCCGG---GACCC 2454
DB 2475 CTGCCAAAGAGCCTCGGAGGACAAAGTCTCTATCCCGCCGACCCCGGGAGGACCC 2534
QY 2455 TCTGTCTTGGCAACAGAGCTCTCAGCTCTCCAAACAGGTGGAGGACCGGAGCAGG 2514
DB 2535 TCTGTCTTGGCAACAGAGCTCTCAGCTCTCCAAACAGGTGGAGGACCGGAGCAGG 2594
QY 2515 CCCCCTGCACCTGGTGGCTGGCAGCTGGGGCGCTGCTCCCGAGCTGCGGCACTGGC 2574
DB 2595 CCCCCTGCACCTGGTGGCTGGCAGCTGGGGCGCTGCTCCCGAGCTGCGGCACTGGC 2654
QY 2575 CTGCAGAAAGCGGGCGGTGAGCTGTGGGGGTCTCCCGGGGACGCGACGCTCCCTGCCI 2634
DB 2655 CTGCAGAAAGCGGGCGGTGAGCTGTGGGGGTCTCCCGGGGACGCGACGCTCCCTGCCI 2714
QY 2635 GATCAGCCCATCGGCCGTGGAGACACAGCTGCGGGGAGCCCTGCCACCTGCGGAG 2694
DB 2715 GATCAGCCCATCGGCCGTGGAGACACAGCTGCGGGGAGCCCTGCCACCTGCGGAG 2774
QY 2695 CTCAGCGCTGTGCTCACCCTGCTCCAAAGAGCTGCGGGCGGGGATTCAGAGCGGCTCAC 2754
DB 2775 CTCAGCGCTGTGCTCACCCTGCTCCAAAGAGCTGCGGGCGGGATTCAGAGCGGCTCAC 2834
QY 2755 AAGTGTGTGGCGCAGGAGCGCGGCTGCTGCGCGGGACCACTGCAACTTGCACCGCAAG 2814
DB 2835 AAGTGTGTGGCGCAGGAGCGCGGCTGCTGCGCGGGACCACTGCAACTTGCACCGCAAG 2894
QY 2815 CCCAGGAGCTGGACTTCTGCTGCTGAGGCGGTGC 2850
DB 2895 CCCAGGAGCTGGACTTCTGCTGCTGAGGCGGTGC 2930
RESULT 5
AAS97182
ID AAS97182 standard; cDNA; 2937 BP.
XX AAS97182;
AC AAS97182;
DT 26-FEB-2002 (first entry)
XX Human metalloprotease partial DNA sequence #11.
DE Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
KW vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;
KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
KW immune-related disease; cardiovascular disease; neuronal disease;
KW migraine; sexual dysfunction; mood disorder; attention disorder;
KW cognition disorder; hypotension; hypertension; psychotic disorder;
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.
XX Homo sapiens.
OS WO200183782-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US14431.
XX PR 04-MAY-2000; 2000US-201879P.
XX PA (SUGEN-) SUGEN INC.
XX PI Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
XX Payne V;
XX WPI; 2002-041502/05.
XX DR P-PSDB; AAU72899.
XX Novel protease polypeptide useful for screening for substances that may

PT be used to treat, e.g., cancers, immune-related diseases,
PT cardiovascular disease, migraine, pain, psychotic and inflammatory
PT disorders -

The invention relates to an isolated, enriched, or purified protease polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to screen for substances (S) that may modulate its activity. Administering S (which modulates protease activity in vitro) may be used to treat a disease or disorder selected from cancers (e.g., of tissues, of blood or hematopoietic origin, of the breast, colon, lung, prostate, cervical, brain, ovarian, bladder or kidney), immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases (e.g., central or peripheral nervous system diseases, migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders and dyskinesias), metabolic disorders and inflammatory disorders. (I) may also be useful as a diagnostic tool for a disease or disorder such as those above. AAS971159-AAS971195 represent human protease coding sequences and primers of the invention.

SQ Sequence 2937 BP; 531 A; 958 C; 936 G; 512 T; 0 other;

Query Match 98.8%; Score 2817.6; DB 24; Length 2937;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2849; Conservative 0; Mismatches 4; Indels 9; Gaps 2;

QY	1	ATGCTTCTGCTGGGATCCTTAACCTTGGCTTTGCGGGCGCAACCGCTTGGAGGCTCTGAG	60
Db	76	ATGCTTCTGCTGGGCATCTTAACCTTGGCTTTGCGGGCGCAACCGCTTGGAGGCTCTGAG	135
QY	61	CCAGAGCGGGAGGTAGTCGTTCCCATCCGACTGGACCGCGGACATTAAACGGCCGCGGTAC	120
Db	136	CCAGAGCGGGAGGTAGTCGTTCCCATCCGACTGGACCGCGGACATTAAACGGCCGCGGTAC	195
QY	121	TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGGACTCATTTTTCAGATCACAGCATTT	180
Db	196	TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGGACTCATTTTTCAGATCACAGCATTT	255
QY	181	CAGGAGGACTTTTACCTACACCTAGCCCGGATGCTCAGTTCCTTGGCTCCCGCTTCTCC	240
Db	256	CAGGAGGACTTTTACCTACACCTAGCCCGGATGCTCAGTTCCTTGGCTCCCGCTTCTCC	315
QY	241	ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACTGCGACGC	300
Db	316	ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACTGCGACGC	375
QY	301	TGCTTCTATTCTGGGGACGTGAACGCCGAGCGGACTCGTTGCTGTCTGTAGGCTGTGC	360
Db	376	TGCTTCTATTCTGGGGACGTGAACGCCGAGCGGACTCGTTGCTGTCTGTAGGCTGTGC	435
QY	361	GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCAATTAAGCCCGCTGCC	420
Db	436	GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCAATTAAGCCCGCTGCC	495
QY	421	AATGCTAGCGCGCGGGCGAGCGACACAGCCAGGGCGACACACTTCTCCAGCGCGG	480
Db	496	AATGCTAGCGCGCGGGCGAGCGACACAGCCAGGGCGACACACTTCTCCAGCGCGG	555
QY	481	GGTGTTCGGCGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG	540
Db	556	GGTGTTCGGCGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG	615
QY	541	AACCCCGCATCTACGGGCCCTTGACCTTACAAGCCGAGCGCGGGCGGCTTCGGGGAG	600
Db	616	AACCCCGCATCTACGGGCCCTTGACCTTACAAGCCGAGCGCGGGCGGCTTCGGGGAG	675
QY	601	AGTCGTAGCGGGCGAGGCTCGGGCGCGCAAGCGTTTTCGTCTATTCGCGCGGTACGTG	660
Db	676	AGTCGTAGCGGGCGAGGCTCGGGCGCGCAAGCGTTTTCGTCTATTCGCGCGGTACGTG	735
QY	661	GAGACGCTGTGTCCGGGACGAGTCAATGTTCAGATTTCACGCGCGGACCTTGGAACT	720

Db	1816	GAGGAGCAGTGTGAGGCTTTCAACGGCTTAAACACACAGCAACCAACCGGCCTCACTCTCGCC	1875
Qy	1801	GTGGCATGGGTGCCCAAGTACTTCGGCGCTGTCTCCCGGGACAAGTGCAAGCTCATCTGC	1860
Db	1876	GTGSCATGGTGGCCCAAGTACTTCGGCGCTGTCTCCCGGGACAAGTGCAAGCTCATCTGC	1935
Qy	1861	CGAGCCAATGGCACTGGCTACTTCTATGTCTGGCACCCCAA--GGTGGTGGACGGCAGC	1917
Db	1936	CGAGCCAATGGCACTGGCTACTTCTATGTCTGGCACCCCAAAGTGGTGGTGGACGGCAGC	1995
Qy	1918	CTGTGCTCTCTGACTCCACCTCCGCTGTGTCTCAAGGCAAGTGCAATCAAGGCTGGCTGT	1977
Db	1996	CTGTGCTCTCTGACTCCACCTCCGCTGTGTCTCAAGGCAAGTGCAATCAAGGCTGGCTGT	2055
Qy	1978	GATGGGAACCTGGGCTCCAAAGAAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAT	2037
Db	2056	GATGGGAACCTGGGCTCCAAAGAAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAT	2115
Qy	2038	AAGAGCTGCAAGAAGTGACTGGACTCTTCACCAAGCCCATGCAATGCTACAATTTCGTG	2099
Db	2116	AAGAGCTGCAAGAAGTGACTGGACTCTTCACCAAGCCCATGCAATGCTACAATTTCGTG	2175
Qy	2098	GTGGCCATCCCGCAGCGCCCTCAAGCATCGACATTCGCCAGCCGCTTACAAGGGCTG	2157
Db	2176	GTGGCCATCCCGCAGCGCCCTCAAGCATCGACATTCGCCAGCCGCTTACAAGGGCTG	2235
Qy	2158	ATCGGGGATGACAACCTACCTGGCTGTGAAGAACACGCAAGCAAGTACCTTCCTCAACGGG	2217
Db	2236	ATCGGGGATGACAACCTACCTGGCTGTGAAGAACACGCAAGCAAGTACCTTCCTCAACGGG	2295
Qy	2218	CATTTCGTGTGTGCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTCTCGCGGTAC	2277
Db	2296	CATTTCGTGTGTGCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTCTCGCGGTAC	2355
Qy	2278	AGCGGCACGGGCACACGGTGGAGAGCCTGTGAGGTTTCCCGGCCCATCTCTCGAGCCGCTG	2337
Db	2356	AGCGGCACGGGCACACGGTGGAGAGCCTGTGAGGTTTCCCGGCCCATCTCTCGAGCCGCTG	2415
Qy	2338	ACCGTGGAGTCTCTCTCCGTGGGGAGATGACACCGCCCGGGTCCGCTACTCCTCTCTAT	2397
Db	2416	ACCGTGGAGTCTCTCTCCGTGGGGAGATGACACCGCCCGGGTCCGCTACTCCTCTCTAT	2475
Qy	2398	CTGCCAAAGAGCCTCGGAGGACAAAGTCTCTCATCTCCAAAGGACCCCGG-----GGA	2451
Db	2476	CTGCCAAAGAGCCTCGGAGGACAAAGTCTCTCATCTCCCGCCACCCCGGAGGAGGA	2535
Qy	2452	CCCTCTCTCTTGACAAACAGCGTCTCAGCCTCTCCACCAAGTGGAGCAGCGGACGAC	2511
Db	2536	CCCTCTCTCTTGACAAACAGCGTCTCAGCCTCTCCACCAAGTGGAGCAGCGGACGAC	2595
Qy	2512	AGGCCCCCTCGACGCTGGGTGGCTGGCAGCTGGGGCCGCTGCTCCGCGAGCTGGCGCAGT	2571
Db	2596	AGGCCCCCTCGACGCTGGGTGGCTGGCAGCTGGGGCCGCTGCTCCGCGAGCTGGCGCAGT	2655
Qy	2572	GGCTGACAGAAAGGGGGGTGGACTGTCCGGGCTCCGGGGGACGACCGTCCCTGCC	2631
Db	2656	GGCTGACAGAAAGGGGGGTGGACTGTCCGGGCTCCGGGGGACGACCGTCCCTGCC	2715
Qy	2632	TGTGATCAGCCCATCGGCCGTGGAGACACAAGCCTGCGGGAGCCCTGCCCCACCTGG	2691
Db	2716	TGTGATCAGCCCATCGGCCGTGGAGACACAAGCCTGCGGGAGCCCTGCCCCACCTGG	2775
Qy	2692	GAGCTCAGCGCCTGGTCAACCTGCTCCAAAGAGTGGGGCCGGGATTTTCAGAGGGCTCA	2751
Db	2776	GAGCTCAGCGCCTGGTCAACCTGCTCCAAAGAGTGGGGCCGGGATTTTCAGAGGGCTCA	2835
Qy	2752	CTCAAGTGTGTGGGCCACGAGGCGCGGTGTGTGGCCCGGGACAGTGCACCTTGACCGC	2811
Db	2836	CTCAAGTGTGTGGGCCACGAGGCGCGGTGTGTGGCCCGGGACAGTGCACCTTGACCGC	2895
Qy	2812	AAGCCCCAGGAGCTGGACTTCTCGCTCTGAGGCGGTGTGTA	2853
Db	2896	AAGCCCCAGGAGCTGGACTTCTCGCTCTGAGGCGGTGTGTA	2937

Db 97 GCGAGGACAGACTCCTTTGCTGCTAAGCCTATGCGGGGTCTCCGCGAGCCTTTGGC 156
QY 385 TACCAGGCGCGAGTATGCTATAGCCCGCTGCCCAATGCTAGCGCGCGCGGGCGCAG 444
Db 157 TACCAAGGTGGGAGTATGCTATTAGCCCTCTGCCCAACACAGCGCGCTGAGGGCGAG 216
QY 445 CGCAACACAGCGGCGCACACTTCTCCAGCGCGGGGTGTTCGGGCGCGCCCTTCGCGA 504
Db 217 CGTATAGCGAGGCGCACACTTCTCCAGCGCGGGGTCTCCGCTAGGCGCTTCGCGA 276
QY 505 GACCCCACTCTCGCTGCGGGGTGGCTTCGGCTGGAAACCCCGGCATCTACGCGGCGCTG 564
Db 277 GACCCCTACCTCTCGCTGCGGGGTGGCTTCGGCTGGAAACCCCGGCATCTACGCGCTG 336
QY 565 GACCCCTACAAGCGCGGCGGCGCTTCGGGAGAGTCTAGCCGCGCGAGCTCGGG 624
Db 337 GACCCCTTAAACACAGCGGCGGCGGTGGCGGAAGCCCAACACCGCGCGAGCTCTGGG 396
QY 625 CGCGCAAGCGTTCTGCTATCCCGCGGTACGTGGAGCGCTGGTGGTGGCGGAGCG 684
Db 397 CGCGCAAGCGTTCTGCTATACACGGTACGTGGAGACACTGGTGGTGGCGGAGCGAG 456
QY 685 TCAATGGTCAAGTCCAGCGCGGCGACTCGAACAATTATCTGCTGACCTCTGGCAACG 744
Db 457 TCAATGGTCAAGTCCAGCGCGGCGATTGGAAACATTATCTGCTGACGCTCTGGCCAG 516
QY 745 GCGGCGGACTCTACCGGCATCCAGCATCTCAACCCCATCAACATCTGTTGGTCAAG 804
Db 517 GCGGCGGACTCTACCGGCACCCAGCATCTCAACCTATCAACATCTGTTGGTCAAG 576
QY 805 GTGCTGCTTTAGAGATCGTACCTCCGCGGCCCAAGGTACCGGCAATGGCGCCTGACG 864
Db 577 GTGTTACTTTAGGAGTCTGACACTGGGCGCCAAAGGTACAGGCAACCGCGCCTGACT 636
QY 865 CTGCGCACTCTGTGCTGCGGAGAGAGCTGAACAAGTGAGTGACAAGCAACCCCGAG 924
Db 637 CTGCGCACTCTGTGCTGCGGAGAAAAGTTGAACAAGTGAGCGACAACCCCGAG 696
QY 925 TACTGGGACACTGCCATCCTTCCAGGAGGAGACCTGTGTGGAGCCACACCTGTGAC 984
Db 697 TACTGGGACAGACCATCTCTTCCAGGAGAGGACCTATGCGGGCTTACCACCTGTGAC 756
QY 985 ACCCTGGGATGGCTGATGGTGGTACCATCTGTGACCCCAAGAGAGCTGCTCTCAT 1044
Db 757 ACCTGGGATGGCTGATGGGACCATGTGTGATCCCAAGAGAGCTGCTGTGATC 816
QY 1045 GAGGAGATGGGTTCCATCAGCCTTCACTGCCCCAGAGCTGGGCGCAGTGTTCAC 1104
Db 817 GAGGAGATGGGTTCCGTCGSCCTCACCACCTGCCCCAGCTGGGCGCATGTGTCAAC 876
QY 1105 ATGCCCATGACATGTGAAGTCTGTGAGGAGGTGTTGGAGCTCCGAGCCCAACAC 1164
Db 877 ATGCCCATGACATGTGAAGTGTGTGAGGAGGTGTTGGAGAGCTCAGAGCCCAAC 936
QY 1165 ATGATGTCCCGACCTCATCAGATCGACCTGCGCAACCCCTGGTCAAGCTGCAAGTCT 1224
Db 937 ATGATGTCCGACACTCATCAGATCGACCTGCGCAACCCCTGGTCAAGCTGCAAGTCT 996
QY 1225 GCAATCATCAGGACTTCTGAGAGCGGGACCGGTGACTGCTCTCTGGACCAACCCAGC 1284
Db 997 GCAATCATCAGGACTTCTGAGAGCGGGACCGGTGACTGCTCTCTGGAGAGCTCAGAGCC 1056
QY 1285 AAGCCATCTCCTGCGGAGATCTGCGGGCGGCGACCTACACCTGAGCGCAGAGTGC 1344
Db 1057 AAGCCATCACCCTGCTGAGGACTTGCAGGCGCAAGCTAGAGTTGAGCGCACAGTGC 1116
QY 1345 GAGCTGGCTTTGGCGTGGCTTCAAGCCCTGTCTTACATGACTGACTGACCAAGCTG 1404
Db 1117 GAGCTGGCTTTGGGTTGGCTTCAAGCCCTTCCCATATATGACTGACTGACAAAGCTG 1176
QY 1405 TGTGACCGGGAGCGCAAGGAGAGATGGTGTGCGGACCGCCGACTTCCCTCTGGGCC 1464
Db 1177 TGTGACCGGGAGCGCAAGGAGAGATGGTGTGCGGAGCTGCGGCACTTCCCTCTGGGCA 1236

QY 1465 GATGSCACACAGCTGTGGCGAGGCAAGCTCTGCTCAAGGGGCTCGCTGGAGACAC 1524
Db 1237 GATGSCACACAGCTGTGTGAGGGCAAGTCTGCTCAAGGAGCTCGCTGGAGACAC 1296
QY 1525 AACCTCAACAGCAGAGGTGGATGGTTCTTGGGCCAAATGGGATGCTATGGCCCTGC 1584
Db 1297 AACCCAAACAGTACCGGGTGGACGGCCTTGGGCCAAGTGGGAGCCTTATGGTCCCTGC 1356
QY 1585 TCGCCACATGTGTGGGGGTGCAGCTGGCAGGAGGCA--GTGCACCAACCCCAACC 1642
Db 1357 TCGCCACCTGCGGTGGGGCGCGAGCTGGCCGAGGCAAGTGCAGGACACCCACCC 1416
QY 1643 CTGCCAACGGG--GGCAAGTACTGCGAGGAGTGGGTGAAATACCGATCTTGCATCTG 1701
Db 1417 CTGCCAACGGGCGGAGTACTGCGAGGAGTGGAGTGAATACCGATCTTGCATCTG 1476
QY 1702 GAGCCCTGCCCGAGCTCAGCCTCCGGAAGAGCTTCCGGGA 1742
Db 1477 GAGCCCTGCCCGAGCTCAGCCTTGGCAAGAGCTTCCGGGA 1517

RESULT 7

AAD35570

ID AAD35570 standard; cDNA; 1104 BP.

XX AAD35570;

XX AC 26-JUL-2002 (first entry)

XX Human protease cDNA #3.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;
KW high blood pressure; arthritis; connective tissue disorder; infertility;
KW gene therapy; enzyme; gene; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..1104

FT /tag= a

FT /product= "Human protease #3"

XX WO200226949-A2.

XX PD 04-APR-2002.

XX PF 27-SEP-2001; 2001WO-US30350.

XX PR 29-SEP-2000; 2000US-236689P.

XX (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Hilbun E;

XX WPI; 2002-372123/40.

XX P-PSDB; AAE22542.

XX Novel nucleic acid encoding a human protease, useful as a hybridization
PT probe for screening libraries and assessing gene expression patterns -
XX Disclosure; Page 38-39; 41pp; English.

XX The present sequence is a cDNA encoding novel human protein (NHP).
CC human protease. NHPs share structural similarity with animal proteases
CC particularly zinc metalloproteases. Sequences of the invention are
CC useful in therapeutic, diagnostic and pharmacogenomic applications.
CC NHP polynucleotides are used as hybridisation probes for screening
CC libraries and assessing gene expression patterns. They can also be
CC used for treating related biological disorders such as obesity, high
CC blood pressure, arthritis, connective tissue disorders and infertility.
CC They are also used in gene therapy.

XX

SQ	Sequence 1104 BP; 189 A; 357 C; 345 G; 213 T; 0 other;	
	Query Match 38.2%; Score 1091.2; DB 24; Length 1104;	
	Best Local Similarity 99.7%; Pred. No. 2.3e-218;	
	Matches 1093; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	1 ATGCTTCTGCTGGGATCCTTAACCTGCTTTTCGCGGGCGAACCGCTGGAGCTCTGAG 60	
Db	1 ATGCTTCTGCTGGGATCCTTAACCTGCTTTTCGCGGGCGAACCGCTGGAGCTCTGAG 60	
Qy	61 CCAGAGCGGAGGTAGTCTGCTCCATCCGACTGGAGCCCGACATTAACCGCCGCGCTAC 120	
Db	61 CCAGAGCGGAGGTAGTCTGCTCCATCCGACTGGAGCCCGACATTAACCGCCGCGCTAC 120	
Qy	121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCAGACATTT 180	
Db	121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCAGACATTT 180	
Qy	181 CAGGAGACTTTTACCTTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240	
Db	181 CAGGAGACTTTTACCTTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240	
Qy	241 ACTGAGCATCTGGCGGCTTCCCTCCAGGGGCTCACCGGGGCTTCACACCTGCGACGC 300	
Db	241 ACTGAGCATCTGGCGGCTTCCCTCCAGGGGCTCACCGGGGCTTCACACCTGCGACGC 300	
Qy	301 TGCTTCTATTCTGGGAGCTGAACCGCGGAGCGGACTCGTTGCTGCTGTGAGCCTGTGC 360	
Db	301 TGCTTCTATTCTGGGAGCTGAACCGCGGAGCGGACTCGTTGCTGCTGTGAGCCTGTGC 360	
Qy	361 GGGGGCTCCGCGGAGCTTTTGGCTACCGAGGCGCGGATGATGATCAATAGCCGCTGCC 420	
Db	361 GGGGGCTCCGCGGAGCTTTTGGCTACCGAGGCGCGGATGATGATCAATAGCCGCTGCC 420	
Qy	421 AATGCTAGCGCGCGGGCGGCGAGCGCAACAGCAGGCGGCACACCTTCTCCAGCGCCG 480	
Db	421 AATGCTAGCGCGCGGGCGGCGAGCGCAACAGCAGGCGGCACACCTTCTCCAGCGCCG 480	
Qy	481 GGTGTTCCGGCGGGCTTCCGAGAGCCACCTCTGCTGCGGGGTGGCTCGGCTGG 540	
Db	481 GGTGTTCCGGCGGGCTTCCGAGAGCCACCTCTGCTGCGGGGTGGCTCGGCTGG 540	
Qy	541 AACCCCGCATCTTACGGGCGCTTGACCCCTTACAAGCGCGGGCGGGCTTCGGGGAG 600	
Db	541 AACCCCGCATCTTACGGGCGCTTGACCCCTTACAAGCGCGGGCGGGCTTCGGGGAG 600	
Qy	601 AGTGTAGCGGCGAGCTTGGCGCGCAAGCGTTTCTGCTATCCCGCGGTACGTG 660	
Db	601 AGTGTAGCGGCGAGCTTGGCGCGCAAGCGTTTCTGCTATCCCGCGGTACGTG 660	
Qy	661 GAGAGCTGTGTGCGGAGCTAATGTTCAAGTTCCACGCGCGGACCTTGGAAACAT 720	
Db	661 GAGAGCTGTGTGCGGAGCTAATGTTCAAGTTCCACGCGCGGACCTTGGAAACAT 720	
Qy	721 TATCTGTAGCTGTGTGGCAACCGCGGCGGACTTACCGCATCCACATCTCTCAAC 780	
Db	721 TATCTGTAGCTGTGTGGCAACCGCGGCGGACTTACCGCATCCACATCTCTCAAC 780	
Qy	781 CCCATCAACATCTGTGTGCTCAAGTCTGCTTACAGATCTGACTCCGGGCGCCCAAG 840	
Db	781 CCCATCAACATCTGTGTGCTCAAGTCTGCTTACAGATCTGACTCCGGGCGCCCAAG 840	
Qy	841 GTCACCGCAATGGGCGCTGACGCTCGCAACTTCTGTGCTGGCAGAGAAAGCTGAAC 900	
Db	841 GTCACCGCAATGGGCGCTGACGCTCGCAACTTCTGTGCTGGCAGAGAAAGCTGAAC 900	
Qy	901 AAGTGTAGTACAGACCCCGAGTACTGGGACACTGCCATCTCTTCACGACGACGAC 960	
Db	901 AAGTGTAGTACAGACCCCGAGTACTGGGACACTGCCATCTCTTCACGACGACGAC 960	
Qy	961 CTGTGTGAGGACCACTGTGACACCTTGGCATGCTGATGTGGGTACCATGTGTGAC 1020	
Db	961 CTGTGTGAGGACCACTGTGACACCTTGGCATGCTGATGTGGGTACCATGTGTGAC 1020	

Qy	1021 CCCAAGAGAAGCTCTCTGTCTTATTGAGGAGATGGGCTTCATCAGCTTCCACACTGCC 1080	
Db	1021 CCCAAGAGAAGCTCTCTGTCTTATTGAGGAGATGGGCTTCATCAGCTTCCACACTGCC 1080	
Qy	1081 CACGAGCTGGGCCACG 1096	
Db	1081 CACGAGCTGGGTAAGG 1096	
RESULT 8		
AAA95831		
ID	AAA95831 standard; cDNA; 1143 BP.	
XX	AAA95831;	
AC	AAA95831;	
XX	23-FEB-2001 (first entry)	
DT	Human metalloproteinase ADAMTS-5 cDNA.	
XX	Human; ADAMTS-5; metalloproteinase; ADAM;	
DE	a disintegrin and metalloproteinase domain; thrombospondin domain;	
XX	vacine; nootropic; neuroprotective; antiparkinsonian;	
KW	cerebroprotective; cytosolic; antiarthritic; immunosuppressive;	
KW	Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;	
KW	autoimmune disease; brain tumour; brain injury; ss.	
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1143
FT		/*tag= a
FT		/partial
FT		/product= "ADAMTS-5"
XX		
XX	WO200053774-A2.	
PN		
XX	14-SEP-2000.	
PD		
XX	08-MAR-2000; 2000WO-US06237.	
PF		
XX	08-MAR-1999; 99US-0264585.	
PR		
XX	(NEUR-) NEUROCRINE BIOSCIENCES INC.	
PA		
XX	Kelner GS, Clark M, Maki RA;	
PI		
XX	WPI; 2000-594326/56.	
DR	P-PSDB; AAB21261.	
DR		
XX	Polynucleotide encoding novel members of a disintegrin,	
PT	metalloproteinase and thrombospondin domain protein family used to	
PT	prevent and treat Alzheimer's disease, cancer and autoimmune diseases -	
PT	Claim 2; Fig 22; 129pp; English.	
XX		
CC	The present sequence encodes human metalloproteinase ADAMTS-5. The	
CC	ADAMTS family of proteins is closely related to the ADAM (A Disintegrin	
CC	and Metalloproteinase Domain) family. Members of the ADAMTS family	
CC	contain a thrombospondin domain in addition to the disintegrin and	
CC	metalloproteinase domains found in the ADAMTS. ADAMTS polypeptides are	
CC	useful for the manufacture of medicaments for treating conditions	
CC	associated with neuroinflammation and/or neurodegeneration, such as	
CC	Alzheimer's disease, Parkinson's disease and stroke. They are also	
CC	useful for treating conditions associated with cell proliferation, cell	
CC	migration, inflammation and/or angiogenesis, such as cancer, arthritis	
CC	and autoimmune diseases. They can be used to treat patients afflicted	
CC	with an invasive tumour, a brain tumour or brain injury.	
XX		
SQ	Sequence 1143 BP; 240 A; 362 C; 326 G; 209 T; 6 other;	
	Query Match 36.6%; Score 1043.6; DB 21; Length 1143;	
	Best Local Similarity 96.5%; Pred. No. 2e-208;	

Matches 1084; Conservative 2; Mismatches 34; Indels 3; Gaps 2;

```
QY 738 GCGAAGCGGGCGGAGCTCTACCGCATCCCGAGCATCTCAACCCCATCAACATCGTTGT 797
Db 21 GCGCGCGCGGGGAGGTATCTTTAAGATCCCGAGCATCCCGAGCATCAACCCCATCAACATCGTTGT 80
QY 798 GGTCAAGTGTCTCTTTAGAGTCTGTGACTCCGCGCCCAAGTCAACCGCAATCGCGC 857
Db 81 GGTCAAGTGTCTCTTTAGAGTCTGTGACTCCGCGCCCAAGTCAACCGCAATCGCGC 140
QY 858 CTGACGCTGCGCAACTTCTGTGCTGGCAGAAGAGTGAACAAAGTGAAGTGAACAAGCA 917
Db 141 CCTGACGCTGCGCAACTTCTGTGCTGGCAGAAGAGTGAACAAAGTGAAGTGAACAAGCA 200
QY 918 CCCCAGTACTGGGACACTGCCATCTCTTACACAGGAGGACCTGTGTGAGCCACAC 977
Db 201 CCCCAGTACTGGGACACTGCCATCTCTTACACAGGAGGACCTGTGTGAGCCACAC 260
QY 978 CTGTGACACCTTGGCATGGCTGTGATGGGTACCATGTGTGACCCCAAGAGAACTGCTC 1037
Db 261 CTGTGACACCTTGGCATGGCTGTGATGGGTACCATGTGTGACCCCAAGAGAACTGCTC 320
QY 1038 TGTCAATTGAGGACGATGGCTTCCATCAGCCTTCAACCACTCCCAACGAGTGGGCCACGT 1097
Db 321 TGTCAATTGAGGACGATGGCTTCCATCAGCCTTCAACCACTCCCAACGAGTGGGCCACGT 380
QY 1098 GTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTGTGTGGAAAGTCCGAGC 1157
Db 381 GTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTGTGTGGAAAGTCCGAGC 440
QY 1158 CAACACATGATGTCCCGACCTTCATCCAGATCGACGCTGCCAACCCCTGGTCAGCCTG 1217
Db 441 CAACACATGATGTCCCGACCTTCATCCAGATCGACGCTGCCAACCCCTGGTCAGCCTG 500
QY 1218 CAGTCTGCCATCATACCGACTTCTTGACAGCGGCGACGGTGAAGTCTCTCTGGACCA 1277
Db 501 CAGTCTGCCATCATACCGACTTCTTGACAGCGGCGACGGTGAAGTCTCTCTGGACCA 560
QY 1278 ACCAGCAAGCCCATCTCCCTGCCGAGGAGTCTGCCGGGCGCCAGCTACACCTTGAGCCA 1337
Db 561 ACCAGCAAGCCCATCTCCCTGCCGAGGAGTCTGCCGGGCGCCAGCTACACCTTGAGCCA 620
QY 1338 CGATGCGAGTGGCTTTTGGCGTGGCTCCAGCCCTGCTTACATGAGTACTGCAC 1397
Db 621 CGATGCGAGTGGCTTTTGGCGTGGCTCCAGCCCTGCTTACATGAGTACTGCAC 680
QY 1398 CAAGCTGTGTGACACCGGAGGCAAGGGACAGATGCTGTGCAGACCGCCACTTCCC 1457
Db 681 CAAGCTGTGTGACACCGGAGGCAAGGGACAGATGCTGTGCAGACCGCCACTTCCC 740
QY 1458 CTGGGCGGATGGCACCAGCTGTGGCGAGGGAAGTCTGCTCTCAAAGGGCCCTGCGTGA 1517
Db 741 CTGGGCGGATGGCACCAGTGTGGCGAGGGAAGTCTGCTCTCAAAGGGCCCTGCGTGA 800
QY 1518 GAGACAACTCAACAGACAGGTTGATGTTCTGCGCCCAATGGATCCCTATGG 1577
Db 801 AARACAACTCAACAGACAGGTTGATGTTCTGCGCCCAATGGATCCCTATGG 860
QY 1578 CCCCTGCTCGCGCACATGTGTGGGGCGGTGCAGCTGGCCAGGAGGAG-TGCACCAAC 1636
Db 861 CCCCTGCTCGCGCACATGTGTGGGGCGGTGCAGCTGGCCAGGAGGAGTGCACCAAC 920
QY 1637 CCACCCC--TGCCAACGGGGCAAGTACTCGAGGGAGTGAAGTAAATACCGATCCTG 1694
Db 921 CCANCCCTCGCAACNGGGGGCAAGTACTCGAGGGAGTGAAGTAAATACCGATCCTG 980
QY 1695 CAATCTCGGCGCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGA 1754
Db 981 CAACCTGGAGCGCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGA 1040
QY 1755 GGCTTTTCAACGGGTACAAACACAGCACCAACCGGCTCAGTCTCGCGTGGCATGGGTGCC 1814
Db 1041 GGCTTTTCAACGGGTACAAACACAGCACCAACCGGCTCAGTCTCGCGTGGCATGGGTGCC 1100
```

```
QY 1815 CAAGTACTCGGGGTGTCTCCCGGGACAAAGTGCAAGCTCATC 1857
Db 1101 CAAGTACTCGGGGTGTCTCCCGGTGACAAAGTGAAGCTCATC 1143

RESULT 9
AAD35568
ID AAD35568 standard; cDNA; 966 BP.
XX
AC AAD35568;
XX
26-JUL-2002 (first entry)
XX
Human protease cDNA #1.
XX
Human; novel human protein; NHP; protease; biological disorder; obesity;
high blood pressure; arthritis; connective tissue disorder; infertility;
gene therapy; enzyme; gene; ss.
XX
Homo sapiens.
XX
Key Location/Qualifiers
FT CDS 1..966
FT /tag= a
FT /product= "Human protease #1"
XX
WO200226949-A2.
XX
04-APR-2002.
XX
27-SEP-2001; 2001WO-US30350.
XX
29-SEP-2000; 2000US-236689P.
XX
(LEXI-) LEXICON GENETICS INC.
XX
Fridde CJ, Hilbun E;
PI WPI; 2002-372123/40.
DR P-PSDB; AAE22540.
XX
Novel nucleic acid encoding a human protease, useful as a hybridization
probe for screening libraries and assessing gene expression patterns -
Disclosure; Page 34; 4lpp; English.
XX
The present sequence is a cDNA encoding novel human protein: (NHP),
human protease. NHPs share structural similarity with animal proteases
particularly zinc metalloproteases. Sequences of the invention are
useful in therapeutic, diagnostic and pharmacogenomic applications.
NHP polynucleotides are used as hybridisation probes for screening
libraries and assessing gene expression patterns. They can also be
used for treating related biological disorders such as obesity, high
blood pressure, arthritis, connective tissue disorders and infertility.
They are also used in gene therapy.
XX
Sequence 966 BP; 161 A; 316 C; 305 G; 184 T; 0 other;

Query Match 33.6%; Score 959; DB 24; Length 966;
Best Local Similarity 100.0%; Pred. No. 8.8e-191;
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTCGGCGATCCTTACCCCTGGCTTCGCGGGCGAACCCTGGAGGCTTCAG 60
Db 1 ATGCTTCTGCTCGGCGATCCTTACCCCTGGCTTCGCGGGCGAACCCTGGAGGCTTCAG 60
QY 61 CCAGAGCGGGAGGTAGTCTGCCATCCGACTGGACCGGACATTAACGGCGCCCTAC 120
Db 61 CCAGAGCGGGAGGTAGTCTGCCATCCGACTGGACCGGACATTAACGGCGCCCTAC 120
QY 121 TACTGCGGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180
Db 121 TACTGCGGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180
```

Db	121	TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGGACTCAATTTTTCAGATCACAGCATTT	180
Qy	181	CAGGAGGACTTTTACCTACACCTGACGCGCGGATGCTCAGTTC	240
Db	181	CAGGAGGACTTTTACCTACACCTGACGCGCGGATGCTCAGTTC	240
Qy	241	ACTGAGCATCTGGGCTCCCCCTCCAGGGGCTCACGGGGGCTCTTCAGACTGGGACGC	300
Db	241	ACTGAGCATCTGGGCTCCCCCTCCAGGGGCTCACGGGGGCTCTTCAGACTGGGACGC	300
Qy	301	TGCTTCTATCTGGGACGTGAACCGCGAGCCGGACTCGTTGCTGCTGTGAGCCCTGTC	360
Db	301	TGCTTCTATCTGGGACGTGAACCGCGAGCCGGACTCGTTGCTGCTGTGAGCCCTGTC	360
Qy	361	GGGGGCTCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCAATAGCCCGCTGCC	420
Db	361	GGGGGCTCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCAATAGCCCGCTGCC	420
Qy	421	AATGCTAGCGCGCGCGCGGCGAGCGCAACAGCGAGGGCGACACCTTCTCAGCGCCGG	480
Db	421	AATGCTAGCGCGCGCGCGGCGAGCGCAACAGCGAGGGCGACACCTTCTCAGCGCCGG	480
Qy	481	GGTGTTCGGGCGGGCTTCCGAGACCCACCTCTCGCTGCGGGGTGGCTTCGGGCTGG	540
Db	481	GGTGTTCGGGCGGGCTTCCGAGACCCACCTCTCGCTGCGGGGTGGCTTCGGGCTGG	540
Qy	541	AACCCGCCATCTACGGGCCCTGGACCCCTTACAAGCGCGGCGGGCTTCGGGGAG	600
Db	541	AACCCGCCATCTACGGGCCCTGGACCCCTTACAAGCGCGGCGGGCTTCGGGGAG	600
Qy	601	AGTCGTAGCGCGCGCAGGTCGTGGCGCGCAACAGCGTTTGGTCTATCCCGCGGTACGTG	660
Db	601	AGTCGTAGCGCGCGCAGGTCGTGGCGCGCAACAGCGTTTGGTCTATCCCGCGGTACGTG	660
Qy	661	GAGACGTGCTGTCGGGACGAGTCAATGGTCAAGTTCACGGCGCGGACCTGGAACAT	720
Db	661	GAGACGTGCTGTCGGGACGAGTCAATGGTCAAGTTCACGGCGCGGACCTGGAACAT	720
Qy	721	TATCTGCTGACGCTGTGGCAAGGGGGCGGCGACTCTACCGCATCCAGCATCTCTCAAC	780
Db	721	TATCTGCTGACGCTGTGGCAAGGGGGCGGCGACTCTACCGCATCCAGCATCTCTCAAC	780
Qy	781	CCCATCAACATCGTTGTGGTCAAGGTGCTGCTTTAGACATCGTGACTCCGGGCCCAAG	840
Db	781	CCCATCAACATCGTTGTGGTCAAGGTGCTGCTTTAGACATCGTGACTCCGGGCCCAAG	840
Qy	841	GTCACCGCAATCGGGCCCTGACGCTCGCAACTTCTGTGCTGGCAGAGAAAGCTGAAC	900
Db	841	GTCACCGCAATCGGGCCCTGACGCTCGCAACTTCTGTGCTGGCAGAGAAAGCTGAAC	900
Qy	901	AAAGTGAGTCACAGCACCCCGAGTACTGGGACACTGCCATCCTCTCACCAGGCAGGA	959
Db	901	AAAGTGAGTCACAGCACCCCGAGTACTGGGACACTGCCATCCTCTCACCAGGCAGGA	959

RESULT 10

RESOLI TO
AAH20226

ID AAH20226 standard; DNA; 2670 BP.

XX
XX

AC ААН20226;

XX
XX

DT 07-AUG-2001 (first entry)

XX DE Human ADAM-type metalloprotease-related probe, SEQ ID NO:3.

XX
THE UNIVERSITY OF TEXAS AT AUSTIN

KW Human; MDTs4; ADAM-type metalloprotease; drug screening;

KW A Disintegrin And Metalloprotease; cancer; arthritis; pr

XX

OS Homo sapiens.

XX

PN JP2001017183-A.

XX

XX	09-JUL-1999;	99JP-0196584.	
XX	XX		
XX	09-JUL-1999;	99JP-0196584.	
XX	XX		
XX	(YAMA)	YAMANOUCHI PHARM CO LTD.	
XX	XX		
XX	WPI;	2001-275950/29.	
XX	XX		
XX	A new metal	protease and its preparation for use as an anti-cancer and	
XX	anti-arthritis	therapeutic	
XX	Example 1;	Page 11-12; 22pp; Japanese.	
XX	XX		
XX	The invention	relates to the novel human ADAM (A Disintegrin And	
CC	Metalloprotease)-type metalloproteases	MDTS4 (AAB73549) and MDTS5	
CC	(AAB73550). The metalloproteases	can be used for the treatment of	
CC	cancers and arthritis. The invention	also relates to the genes encoding	
CC	MDTS4 and MDTS5, vectors and host	cells containing the MDTS4 or MDTS5	
CC	genes, the recombinant production	of MDTS4 and MDTS5, and antibody	
CC	specific for MDTS4 or MDTS5, and	methods of screening for compounds	
CC	which modulate the activity of	MDTS4 and/or MDTS5. The present	
CC	sequence represents a human	ADAM-type metalloprotease-related cDNA	
CC	sequence, used in an exemplification	as a probe for screening human cDNA	
CC	libraries in order to isolate	MDTS4 cDNA.	
XX	XX		
XX	Sequence	2670 BP; 511 A; 853 C; 867 G; 439 T; 0 other;	
XX	XX		

Db	754	CGAATCTACAGCACCCAGCATCAAGAATTCACAACTGATGGTGTGTTAAAGTGCCTG	813
QY	811	CTTCTTAGAGATCGTACTTCGGGCCCCAAGGTACACGGCAATCGGCCCTTGACGCTGC	870
Db	814	ATCGTAGAAGATGAATAATGGGCCAGAGGTGCCGACAATGGGGCTTACACTGCCT	873
QY	871	AACCTCTGTGCTGGCAGAGAGACTGAACAAGTGTAGTACAGAGACACCAGTACTGG	930
Db	874	AACCTTGTCAACTGCGACGCGCTTTCAACGAGCCAGCAGCCGACCCAGAGCACTAC	933
QY	931	GACACTGCCATCCTCTTACCAGGCAGGACCTGTGTGG---AGCCACCACTGTGACACC	987
Db	934	GACACGGCATCTCTCTTACCAGACAGAACTCTGTGGCAGGAGGGCTGTGTGACACC	993
QY	988	CTGGCATGGCTGATGTGGGTACCATAGCTTACCACCTGCCACAGCTGGGCCACGTGTTCAACATG	1047
Db	994	CTGGGTGTGGCAGACATGGGACCAATTTGTACCCCAACAAGAGCTGCTCGTGTATCGAG	1053
QY	1048	GAGATGGCTTCCATAGCCTTACCACCTGCCACAGCTGGGCCACGTGTTTCAACATG	1107
Db	1054	GATGAGGGCTCCAGCGGCCACACCTTGGCCCATGAATTAGGGCAGTCTCTCAGCATG	1113
QY	1108	CCCATGCAATGTGAAGTCTGTGAGGAGTGTTTGGAGCTCCGAGCTCAACACACATG	1167
Db	1114	CCCCAGCAGCTCCAGGCCCTGACACAGCTCTTGGGGCCCATGGCAAGCACAGCTG	1173
QY	1168	ATGTCCTCCGACCCCTCATTCAGATTCAGCGTGCACACCTCTGTGTGAGCTTGCATGTGCC	1227
Db	1174	ATGGCACCCGTGTGTGCTCCACCTGAACAGAGCGTGGCCTGGTCCCTTCGACGGCCATG	1233
QY	1228	ATCATCACGACTCTCTGTGACAGCGGACAGGTGACTTCCCTCCCTGGACCAACCCAGCAAG	1287
Db	1234	TATCTCACAGACTTCTTGGACGGCGGGCACGGAGACTGTCTCTTGGATGCCCTGCTGCG	1293
QY	1288	CCCATCTCCTGCCGAGGATTCGCGGGC-----GCCAGCTACACCTTGACCCAGCAG	1341
Db	1294	GCCTTGCCCTCCCAACAGCGCTCCCGGGCGCATGGGCCCTGTACCAGCTGGACCAAG	1353
QY	1342	TGCGAGCTGGCTTTTGGCTGGGCTTCAAGCCCTGTCTTACA-----TGCAGTAC	1392
Db	1354	TGAGGCGAGATTTTGGGCGGATTTCCGCCACTGCCCCACACCTCTGTCTAGAGCTC	1413
QY	1393	TGCACCAAGCTGTGTG---CACCGGGAAGGCCAAGGACAGATGGTGTGCCACACCCG	1449
Db	1414	TGGCGCCAGCTTTGTGTGCCACACTGATGGGCTGAGCCCTGTGCCACAGAAGATGGC	1473
QY	1450	CACCTTCCCTGGCCGATGGCAGCAGCTGTGGCAGGCGCAGCTCTGCCTCAAAGGGCC	1509
Db	1474	AGCCTGCCCTTGGCTCAGCGCAGCCGTGCGGGCCCTGGGCACCTCTGTCTAGAGGACG	1533
QY	1510	TG-----CGTGGAGAGACACAACCTCAACAGACAGGTTGGATGTCTCTGGGCCAAA	1563
Db	1534	TGCTACCTGAGSAGGAAGTGGAGAGGCCCAAGCCCTGTAGATGAGGCTGGGCCACCG	1593
QY	1564	TGGATTCCTATGGCCCTGTCTCGGCACATGTGTGTGGGCGGTGCAGCTTGGCCAGGAG	1623
Db	1594	TGGGGACCTTGGGGAAGATTTCTCGGACCTGTGGAGGAGGTACAGTTTTCACACCGT	1653
QY	1624	CAGTGCACCAACCCCAACCTTCCAAACGGGGCAAGTACTTGCAGGAGTTCAGGCTGAAA	1683
Db	1654	GAGTGCAAGGACCCGAGCCTCAGATGGAGGAATACTGCTTGGGTTCGAGAGCCAG	1713
QY	1684	TACGATTCCTGCAATCTGAGGCCCTGCCACAGCTCAGCCTCCGGAAAGAGCTTCCGGGAG	1743
Db	1714	TACAGTCAATGCCACAGGAGGAATGCCCC-----CCTGACGGGAAAGTTTCAGGGAG	1767
QY	1744	GAGCAGTGTGAGCTTTCAACGCTTACACACAGCACACCAACCGCTCACTCTCGCGTG	1803
Db	1768	CAGCAGTGTGAGGAAGTATATGCTTCAATTTACATGTGACATGGACGGGAATCT---CCTG	1824
QY	1804	GCATGGTGGCCCAAGTACTCGGCGGTGTCTCCCGGGGACAAGTCAAGCTCATCTGCCGA	1863

Db	1825	CAGTGGGTCCCAAGTATCTGGGTGTCTCCCGCGGACCCCTGCAGTGTCTTCTGTGCGGA	1888
Qy	1864	GCCAATGGCACTGGCTACTTCTATGTGTGGGCACCCCAAGGTGGTGAGCGGACACCTGTGC	1923
Db	1885	GCCTGGGGAGGAGCGAGTGTAAAGTGTTCGAGGCCAAAGTGTATGTATGGCACCTCTTGT	1944
Qy	1924	TCCTCTGACTCCACCTCCGCTCTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGTATGGG	1983
Db	1945	GGGCGGAGAAACACTGGCCATCTGTGTCCGTGGCCAGTGTGTCAAGGCGCGCTGTGACCAT	2004
Qy	1984	AACCTGGGCTCCCAAGAGAAGATTCGACAAAGTGTGGGGTGTGTGGGGGAGACAATAAGAGC	2043
Db	2005	GTGGTGGACTCGCTCGGAAGCTGGACAAAATCGGGGTGTGTGGGGGCAAAAGGCAACTCC	2064
Qy	2044	TGCNAGAGGTGACTGGACTCTTCACCAAGCCCCATGCATGGCTACAAATTCGTGTGGCC	2103
Db	2065	TGCAGGAAGTCTCCGGGTCCCTTCACCCCCCAATTTGGCTACAAATGTGATTCACCT	2124
Qy	2104	ATCCCCCGAGCGCCTCAAGCATCGACATCCGCCAGCGGGTTTACAAAGGCTGATCGGG	2163
Db	2125	ATCCCCACTGGTGCACCTAATATTGACGTGAAGCAGCGGAGCCACCGGGTGTGCAGAAC	2184
Qy	2164	GATGACAACCTACTGGCTCTGAAGACACGCCAAGCAGATGACTCTCTCAACGGGCAATTC	2223
Db	2185	GATGGGAACCTACCTGGCGCTGGAAGACGGCTGATGGGSCAGTACTCTCTCAACGGCAACTG	2244
Qy	2224	GTGCTGTCCGGGTGGAGCGGGACCTGGTGGTGAAGGGCAGTCTCTCGGTTACAGCGGC	2283
Db	2245	GCCATCTCTGCCATAGACGAGGACATCTTGTGTGAAGGGACCATCTCTGAAGTACAGCGC	2304
Qy	2284	ACGGGCACAGCGGTGGAGAGCCTTCAGGCTTCCCGGCCCATCTCTGGAGCCGCTGACCGTG	2343
Db	2305	TCCATCGCCACCTTGGAGCGCTTCAGAGCTTCCTGGCGCCCTGCCAGAGCCTCTGACAGTG	2364
Qy	2344	GAGTCTCTCTCCGT--GGGGAAGATGACACCGCCCCCGGTCCGTACTCTCTTATCTG	2400
Db	2365	CAGTCTCTGACGTCCCTGGCGAGGCTTCTCCCGCCCAAAAGTCAAAATACACCTCTCTTGT	2424
Qy	2401	CCCAAGA 2408	
Db	2425	CCTAATGA 2432	
RESULT 11			
AAF82166			
ID	AAF82166 standard; cDNA; 2670 BP.		
XX	AAF82166;		
XX			
DT	02-JUL-2001 (first entry)		
XX	Human ADAM type metal protease MDR53 encoding cDNA SEQ ID NO:21.		
DE			
XX	Human; a disintegrin and metalloprotease type metal protease: MDR53;		
KW	MDR53; MDR53; ADAM type metal protease; cytosolic; antithrombotic;		
KW	cancer; arthritis; arthrosis deformans; ss.		
XX			
OS	Homo sapiens.		
XX			
PH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	1..2670		
FT	/*tag= a		
FT	/product= "MDR53"		
FT	/note= "a disintegrin and metalloprotease (ADAM)		
FT	type metal protease"		
XX			
PN	JP2001008687-A.		
XX			
PD	16-JAN-2001.		
XX			
PF	25-JUN-1999; 99JP-0180973.		
XX			
PR	25-JUN-1999; 99JP-0180973.		

Db 817 ATCTAGAAATGAAAAATGGGCCAGAGGTGTCGACAAATGGGGGCTTACACTGCGT 876
Qy 871 AACTTCGTGCGCTGGCAGAGAGCTGAACAAAGTGAAGTACACAGCACCAGTACTGG 930
Db 877 AACTTCGTGCACTGGCAGCGGCTTTCAACAGCCAGAGCGCCAGCCAGAGCACTAC 936
Qy 931 GACACTGCCATCTCTTACACAGGAGGACCTGTGTGG--AGCCACCACCTGTGACACC 987
Db 937 GACAGGCCATCTCTACACAGAGAACTCTGTGGCAGGAGGGCTGTGTGACACC 996
Qy 988 CTGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGTCTGTGATTGAG 1047
Db 997 CTGGGTGTGCAGACATCGGGACCAATTTGTGACCCCAACAAAAGCTGCTCGGTGATCGAG 1056
Qy 1048 GACATGGGCTTCCATCAGCTTACCACTGCCAGAGCTGGCCAGCTGGCCAGCTGTTCACATG 1107
Db 1057 GATAGGGGCTCCAGGGGCGCCACCTCGGCCATGAATAGGGCACTGCTTCACATG 1116
Qy 1108 CCCCATGACAAATGTGAAGTCTGTGAGGAGGTGTTGGGAAGCTCCGAGCAACACCATG 1167
Db 1117 CCCACAGCACTCCAGGCCCTGCACAGGCTCTTCGGGCCCATGGCAAGCACACGAG 1176
Qy 1168 ATGTCCCGACCTTCATCAGATGAGACCGTGGCCAAACCCCTGGTACGCTGCACTGTGCC 1227
Db 1177 ATGGCAGCGCTGTGCTCCACCTGAACAGACGCTGCCCCGTGGTCCCGCTGCAGCGGCATG 1236
Qy 1228 ATCATCAGCACTCTTGGACAGGGGACGCTGACCTGCTCTGGACCAACCCAGCAAG 1287
Db 1237 TATCTCAGAGCTCTTGGACGGGGGACGAGAGCTGTCTCTGATGCCCTGGTGGC 1296
Qy 1288 CCCATCTCCCTGCCGAGGATCTGCCGGC-----GCCAGCTACACCTTGACCCAGCAG 1341
Db 1297 GCCCTGCCCTCCACAGGCTCCCGGGCGGATGSCCTGTACAGCTGGACACGAG 1356
Qy 1342 TGCAGCTGGCTTTTGGCGTGGCTTCAAGCCCTGTCTTACA-----TGCAGTAC 1392
Db 1357 TGCAGGAGATCTTTTGGCGCGGATTTCCGCCACTGCCCCACACCTCTGTCTCAGGAGTC 1416
Qy 1393 TGCACCAAGCTGTGGT--CACGGGAAGGCCAGAGGACAGATGGTGTGCCAGACCCGC 1449
Db 1417 TGGCCCAAGCTTTGTGCCACATGATGGGGCTGAGGCCCTGTGCCACAGAAATGGC 1476
Qy 1450 CACTTCCCTGGCGGATGGCAGACAGCTGTGGCAGGGCAAGCTTGCCTCAAGAGGGCC 1509
Db 1477 AGCCTGCCCTGAGCGCACGCCCTGCGGGCTGGGCACTCTGCTCAGAAGGACG 1536
Qy 1510 TG-----CGTGGAGACACAACTCAACAGCACAGGCTGGATGTTCTTGGGCCAA 1563
Db 1537 TGTCTACCTGAGGAGGAAGTGAGAGGCCCAAGGCCCTGTGTAGATGAGGCTGGCACCG 1596
Qy 1564 TGGATCCCTATGCCCTCTGCGCACATGTGTGGGGCGGTGACGCTGGCCAGGAGG 1623
Db 1597 TGGGACCTCGGAGATGTTCTCGACCTGTGGAGAGAGTACAGTTTTCACACCGT 1656
Qy 1624 CAGTGCACAAACCCCTGCCAACAGGGGGCAAGTACTCGAGGAGGTGAGGGTGA 1683
Db 1657 GAGTGAAGAGCCCGAGCCCTCAGATGGAGGAATACTGCTGGGTGGAGGCCAAG 1716
Qy 1684 TACCATCTCCTGATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAG 1743
Db 1717 TACCAGTATGCCACAGGAGGAATGCC-----CCTCAGGGAAAGCTTCAGGGAG 1770
Qy 1744 GAGCAGTGTAGGGTTTCAACGGGTACAAACACAGCACCAGCCGCTCCTCTCGCGGTG 1803
Db 1771 CAGCAGTGTAGAGATATAATGCTCAATTTACACTGACATGGACGGGAATCT--CCTG 1827
Qy 1804 GCATGGGTGCCAAGTACTCTCGGGGTGTCTCCCGGGACAGAGTCAAGCTCATCTGCCGA 1863
Db 1828 CAGTGGGTCCCAAGTATGTGTGGGGTGTCCCGCGGAGCCGCTCAAGTGTGTTCGCCGA 1887
Qy 1864 GCCAATGSCACTGCTACTTCTATGTCTGSCACCCCAAGGTGGTGGACGACGCTGTGC 1923

Db 1888 GCCCGGGGAGGAGGAGTTCAAAAGTGTTCAGGCCAAGGTGATGATGGCACCTGTGT 1947
Qy 1924 TCTCTACTCCACTCGCTCTGTGTCCAGGCAAGTGCATCAAGCTGGCTGTGTATGG 1983
Db 1948 GGGCAGAAACACTGGCCATCTGTGCGTGGCAGTGTGCAAGCGCGGTGTGATAT 2007
Qy 1984 AACTTGGCTCCAGAAGAGATTCGACAAGTGTGGGTGTGGGGGAGACAATAAAGC 2043
Db 2008 GTGTGGTACTCGCTCGGAAGCTGGACAAATGCGGGGTGTGTGGGGCAAGGCAATCC 2067
Qy 2044 TCACAAGAGTGAAGTGGTACTTTCACCAAGCCCATGATGCTGCTCAATTTCTGTGTGGCC 2103
Db 2068 TCAGGAAGGTTCGCGGTCTCCACCCACCAATATGCTACATGACATGTGTCAAC 2127
Qy 2104 ATCCCGCAGCGGCTCAAGCATGACATCCGCGCAGCGCGGTTACAAAGGCTGTATCGG 2163
Db 2128 ATCCAGCTGTGGCACTAATATTGACGTGAAGCAGCGGAGCCACCGGGGTGTGCAAAAC 2187
Qy 2164 GATGACAACTACTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGGCAATTC 2223
Db 2188 GATGGAACTACTGGCTGAAGACGCTGTATGGGCACTACCTGCTCAACGGCAACTG 2247
Qy 2224 GTGTGTGCGGGTGAAGCGGACCTGTGTGAAGGCACTGTCTGCTGCGGTACAGCGC 2283
Db 2248 GCATCTCTGCCATAGACAGACATCTTGTGAAGGGAGCCATCTTGAATGACAGCGC 2307
Qy 2284 ACGGCAAGCGGTGAGAGCTTCAGGCTTCCCGGGCCATCTCTGGAGCCGCTGACCTG 2343
Db 2308 TCCATCGCCACCTGGAGCGCTGCAGAGCTTCGCGCCCTTGCAGAGCCCTTGACAG 2367
Qy 2344 GAGTCTCTCGGT---GGGGAAGATGACAGCGCCCGGGTCCGCTACTCTCTAICTG 2400
Db 2368 CAGCTCTGACAGTCCCTGGGAGGTCTTCCCGCCCAAAAGTCAATACACCTTCTTT 2427
Qy 2401 CCAAGA 2408
Db 2428 CCTATGA 2435

RESULT 13
AAC90058
ID AAC90058 standard; DNA; 3008 BP.
XX
AC AAC90058;
XX
DT 19-MAR-2001 (first entry)
XX
Human METH2 coding sequence.
DE
XX
XX
KW Human; METH2; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control; ss.
XX
OS Homo sapiens.
XX
XX WO200071577-A1.
PN
XX
XX 30-NOV-2000.
XX
XX 25-MAY-2000; 2000WO-US14462.
PF
XX 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.

22-FEB-2000; 2000US-0183792.
 (HUMA-) HUMAN GENOME SCI INC.
 (SMIK-) SMITHKLINE BEECHAM CORP.
 (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 (IRUE-) IREUELA-ARISPE L.
 (HAST-) HASTINGS G A.
 (RUBE-) RUBEN S M.
 (JONA-) JONAK Z L.
 (TRUL-) TRULLI S H.
 (FORN-) FORNWALD J A.
 (TERR-) TERRETT J A.

Iruele-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
 Fornwald JA, Terrett JA;
 WPI: 2001-025136/03.
 N-PSDB; AAC90058.

METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit angiogenesis in the treatment of disorders such as cancer, rheumatoid arthritis and psoriasis -
 Claim 14: Fig 2: 768pp; English.

The present sequence is the coding sequence of human METH2 (ME for metalloprotease and TH for thrombospondin). METH2 can be used for inhibiting angiogenesis in an individual; and for treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculogenesis, granululations, hypertrophic scars, nonunion fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, fibromuscular dysplasia, wound granulation, Crohn's disease or atherosclerosis. METH2 can also be used in birth control. METH2 can also be used in diagnostic methods for the prognosis of cancer.

Sequence 3008 BP: 617 A; 924 C; 948 G; 514 T; 5 other;

Query Match 23.38; Score 664.4; DB 22; Length 3008;
 Best Local Similarity 59.28; Pred. No. 3.4e-129;
 Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

QY 157 CTCATTTTCAGATCACAGCATTTTCAGAGGACTTTTACCTTACACTGACGCGCGATGCT 216
 DB 160 CTCGGCTCCACCTGTCGGCTTCGGCAAGGCTTCGTGTGCGCTTGGCGCCGACGAC 219
 QY 217 CAGTTCTGGCTCCGGCTTCCACTGAGCATCTGGCGGTCCTCCAGGGGCTCACC 276
 DB 220 AGCTTCTTGGCGCGGAGTTCAAGATGAGCGCTCGGGGCTCGGGCGGGG---ACC 276
 QY 277 GGGGCTTTCAGACCTTCGACGCTTCCACTGAGCATCTGGCGGTCCTCCAGGGGCTCACC 336
 DB 277 GGGGCGAGCGGGGCTGCGGCTGTCTTATTTTCGGGACCTGACCGCGAGCGCGAC 336
 QY 337 TCGTTCGCTGTGAGCCTTGGCGGGGCTCGCGGAGCTTTGGCTACCGAGGGGCC 396
 DB 337 TCGTTCGCGGGGTCAGCTGTGCGGGCTGAGCGCTTCTCTGCTGGAGCGGCGAG 396
 QY 397 GAGTATGTCATAGCCGCTGCCAATGCTAGCGCGCGGGGCGGCGAGCGCAACAGCGAC 456
 DB 397 GAGTTCACCATCCAGCGCGAGGGCGGGGGGCTTCCCTGCTGAGCGGCTCAGCGGCTGAG 456
 QY 457 GGGCGACACCTTCTCCAGCGCGGGGTGTTCCGGCGGGCTTCCGGAGACCCCTCT 516
 DB 457 CGCTGGGGTCCCGGAGCGCCCGCTTCCGGAGGACCCGAGTGGGAGGTGGAGACG 516
 QY 517 CGCTCGGGGTG-CCCTCGGGCTGGAAACCGCGCATCTACGGGCGCTGACCTTACAA 575
 DB 517 GGAGAGGCTCAGAGGCGAGGAGAGAGACACCAGGAGACAGCGGGAGGAGACCA 576

QY 576 CCCCGGGGGGGGGCTTCGGGGAGAGTCTGAGCCGGCGGAGG-----TCTGGGCGCGCC 630
 DB 577 GAAGAGGAGGAGAGAGGCGCTAGCGAGCGCCACCGCCCTTGGGGGCGACGAGTAGGACC 636
 QY 631 AAGCGTTTCGTGTCTATCCCGCGTACGTGAGAGCGTGTGTGGTGGCGGAGAGTCAATG 690
 DB 637 AAGCGTTTCGTGTGCTGAGGCGGCTTCTGAGAGCGTGTGTGGTGGCGGAGTGGTCCATG 696
 QY 691 GTCAGTTTCCACGCGCGGACCTTGAACATATATGCTGACGCTGTGTGCAAGCGGGCGG 750
 DB 697 GCTGCTTCTACGGGGCGGACCTGCGAGAACACATCTGACGTTAATGTTCTGTGG-AGCC 756
 QY 751 CGACTCTACGCGCATCCAGCATCTCAACCCATCAACATGTTGTGTGTCAGGTCGCTG 810
 DB 757 CGAATCTACAAGCACCCAGCATCAAGAATTCATCAACCTGATGGTGTGTAAGTCTG 816
 QY 811 CTTCTTAGAGATCGTGACTCCGGGCCCAAGTCAACGGCAATCGGCGCTGACGCTGCGC 870
 DB 817 ATCGTAGAAGATGAATAATGGGCGCCAGAGGTGTCCGACAAATGGGGGCTTTACACTCGGT 876
 QY 871 AACTTCTGTGCTGGCGAGAGAGCTGAACAAAGTGAAGTGAACAGCAACCCCGAGTACTGG 930
 DB 877 AACTTCTGCAACTGGCAGCGGCTTTTCAACCCAGCGGACCCGCGCAGAGCACTAC 936
 QY 931 GACACTGCCATCTCTTCAACGAGGAGGACTGTGTGG---AGCCACCACTGTGACACC 987
 DB 937 GACACGGCCATCTGCTCAACGACAGAACTTCTGTGGCAGGAGGGGTGTGTGACACC 996
 QY 988 CTGGGATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAACTGCTCTGTCTATTGAG 1047
 DB 997 CTGGGTGTGCAGACATCGGGACCAATTTGTGACCCCAACAAAGTGTCTCCGTGATCGAG 1056
 QY 1048 GAGATGGGCTTCCATCAGCTTCAACACTGCCAGCGAGTGGCGGCAAGTGTTCACATG 1107
 DB 1057 GATGAGGGGCTCCAGCGGCGCCACACCTGGCCCATGAACCTAGGAGACGCTCTCAGCATG 1116
 QY 1108 CCCATGACAAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCCACCAACATG 1167
 DB 1117 CCCCAGCAGACTCCAAGCCCTGCACAGGCTCTTCGGGCCCATGGGCAAGCAACAGTG 1176
 QY 1168 ATGTCCCGACCTCATTCAGATCGACCGGTGACCCCTTGCTGACGCTGCACTGTGCGC 1227
 DB 1177 ATGGCACCGCTGTGTCTCCACCTGAACCAAGACGCTGCTGGTCCCTCGAGCGGACATG 1236
 QY 1228 ATCATCAGCAGCTTCTGACAGCGGCGAGTGTGCTCTGTCGACCAACCCAGCAAG 1287
 DB 1237 TATCTCAGAGAGTCTTGGACGCGGGCACGAGACTGTCTCTGGATGCCCTGTGCGG 1296
 QY 1288 CCCATCTCCCTGCGCGGAGTCTGCCGGG------GCCAGCTACACCTGAGCCAGCAG 1341
 DB 1297 GCCCTGCCCTCCCAAGAGGCTCCCGGGCGCATGGCCCTGTACCACTGACCAAGCAG 1356
 QY 1342 TGGGAGCTGGCTTTGGGTGGGCTCCAAGCCCTGTCCTTTACA-----TGCAGTAC 1392
 DB 1357 TGCAGCAGATCTTTGGGCGGATTTCCGCCACTGCCCCAACACCTCTCTCAGGAGCTC 1416
 QY 1393 TGCACCAAGCTGTGGT---CACCGGAGGCGCAAGGACAGATGCTGTGCGAGACCCCG 1449
 DB 1417 TGGCCCGAGTGTGGTGGCCACTGATGGGCTGAGCCCTGTGCCACAGAGATGTCG 1476
 QY 1450 CACTTCCCTGGGCGGATGGCACCACTGTGGCGAGGGCAAGCTGTGCTCAAGGCGGC 1509
 DB 1477 AGCTGCCCTGGGCTGACGCGACGCGGTGCGGGCTTGGGACCTTGTCTCAGAGCAGC 1536
 QY 1510 TG-----CGTGGAGAGACAACTCAACAGCAGAGGTTGGTTCCTGGGCTCAAA 1563
 DB 1537 TGTCTACCTGAGGAGGAGTGGAGAGGCGCAAGCCGCTGTTAGTAGGAGGCTGGGACCG 1596
 QY 1564 TGGGATCCCTATGGCCCTGCTCGGCGACATGTGTGGGGGCGTGCAGCTGCCAGGAGG 1623
 DB 1597 TGGGAGCCCTGGGAGATGTTCTCGGACCTGTGGAGGAGGAGTACAGTTTTCACACCGT 1656
 QY 1624 CAGTGACCAACCCACCCTGCCAACGGGGGCAAGTACTGCGAGGAGTGTAGGCTGAAA 1683

QY	991	GGCATGGCTGATGTGGGTACCATTGTGACCCCAAGAGAGCTGCTCTGHCATTGGAGC	1050
Db	1503	GGAAATGGCAGATGTTTGGAAACGCTATGTGACCCCAAGCAGGAGCTGCTCAGTCATAGAAGAT	1562
QY	1051	GATGGGCTTCCATCAGCCTTCACACACTGCCACAGAGCTGGCCACAGTGTTCACATGCC	1110
Db	1563	GATGGTTTGC AAGCGCCTTCACACAGCCCTGAATTTGGCCATGTGTTTACATPGCG	1622
QY	1111	CATGCAATGTGAAGTCTGTGAGAGGTGTTTGGGAAGCTCGGAGCCACACATGATG	1170
Db	1623	CAGGATGCTTAAGCACTGTGCCAGCTTGAATGGTGAAGTGGCGATCTCATCTGATG	1682
QY	1171	TCCCGACCTCATCCAGATCGACGTGCCAACCCTGTGTGAGCTGCAGTGTGCGCATC	1230
Db	1683	GCTCGATGCTCTCAGCTTAGCATTAGCCAGCCCTTGTTCACCTTGAGTGGCTACATG	1742
QY	1231	ATCACCGACTTCTTGACAGGGCGAGGTGACTGCCTCTTGGACCAACCCAGCAGCC	1290
Db	1743	GTCACTGCTTCTCTAGATAATGGACACGGGAATGTTGTATGGACAAGCCCCAGAATCCA	1802
QY	1291	ATCTCCCTGCCGAGGATCTGCCGGCGCCAGCTACACCTGAGCCAGCAGTGCAGCTG	1350
Db	1803	ATCAAGCTCCCTTCTGATCTTCCCGGTACCTTGTACGATGCCAACCCGAGTGTCA	1862
QY	1351	GCTTTTGGCTGGGCTCCAAAGCCTGTCTTACATGCAGTAC--TGCACCAAGCTGTGG	1407
Db	1863	ACATTCGGAGAGGAATCCAAAGCACTGCCTCTGATGCAGCCAGCATGTACTACCTTGTG	1922
QY	1408	TGCACCGGAAGGCCAAGGACAGATGTGTGCGACGCCCACTTCCTCTGGSCCGAT	1467
Db	1923	TGCACCTGGCACTCCGGTGGCTTACTGTGTGTGCCAAACAACACTTCCCTTGGGAGAT	1982
QY	1468	GGCACCAAGCTGTGGCGAGGGCAAGCTCTGCCTCAAAGGGGCGCTGCGTGGAGACACA	1527
Db	1983	GGCACCAAGCTGTGGAGAAGGAAGTGTGTCTAGTGGCAAGTGCGTGAACAAGACAG	2042
QY	1528	CTCAACAAGCAG--GCTGATGGTTCCTGGGCCCAATGGGATCCCTATGGCCCC	1581
Db	2043	ATGAAGCATTTTGTCTACTCCTCTTCTGAAGCTGGGGACCATGGGACCGTGGGGAG	2102
QY	1582	TGCTCGGCACATGTGTGGGGCGCTGCAGTGTGCCAGAGGCGAGTGCACCAACCCCA	1641
Db	2103	TGCTCAAGAACCTGTGTGTGGAGTTCAATATACAATGAGAGAATGTGACAACCCAG	2162
QY	1642	CTTGCCCAACGGGGCAAGTACTTGCAGGAGGTGAGGTGAAATPACCGATCCTGCAAT	1701
Db	2163	CCAAAGAACGGGAAGTACTGTGAAGCAAAACAGTTCOGCTACAGGTCTCTGTAA	2222
QY	1702	GAGCCCTTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGGCT	1761
Db	2223	GAGGACTGTCCAGACAATAA---CGAATAAGTTCAGAGGAGAGCATGCGAGGCCAC	2279
QY	1762	AACGGGTACAAACACAGCACAAACCGCTTACTCTGCGGTGGCATGGTGTGCCAAGT	1821
Db	2280	AATGAGTGTTC AAGCTTCTCTTGGGAATGAGCCCACTGTAGATGGGACACCCAA	2339
QY	1822	TCCGGGCTGTCTCCCGGGACAAGTGCAGCTCATCTGCCAGAGCAATGGCACTGGCT	1881
Db	2340	GCCGGGCTCTGCAAAGGACAGGTGCAGTCACTTGTGAAGCCAAAGSCATTGGCT	2399
QY	1882	TTCTATGTGTGSCACCCCAAGTGTGTGGACGCAAGCTGTCTCTCTGTACTCCACCT	1941
Db	2400	TTTTTCTCTTTACAGCCCAAGTTGTAGATGGCACTCCCTGTAGTCCGAGACTCTAC	2459
QY	1942	GTCTGTCTCAAGGCAAGTGCAATCAAGGTGGCTGTGATGGAACTGGCTGGCTCCAA	2001
Db	2460	GTCTGTGTGCAAGGCAGTGTGTGAAGCTGGCTGTGATGCCATCATAGACTCCAA	2519
QY	2002	AGATTCGACAAGTGTGGGTGTGTGGGGAGACAATAAGAGCTGCAACAAGGTGACT	2061
Db	2520	AAGTTTCATAAAGTGTGGCTTTGTGGAGAAACGGTCCCATGCAAGAAGATGTCA	2579

Qy	2062	CTCTTACCAGCCCATGCGTACAAATTCCTGGTGGCGCATCCCGCAGGCGCTCA	2122
Db	2580	ATAGTCACCTAGTACAAGAGCCTGGGTATCATGACATTTGTCAAAATTCCTGCTGGAGCCACC	2639
Qy	2122	AGCATCCAGATCCGCCAGGCGGTTACAAAGGGCTGATCGGGGATGACAACTACCTGGCT	2181
Db	2640	AACATTGAGTGAACATCGGAATCAAGGGGGTCCAGAACAAATGGCAGCTTCTGGCT	2699
Qy	2182	CTGAAGAACCCAGCAAGTACCTGCTCAACGGGCAATTCGTGGTGTCTCGCGGTGGAG	2241
Db	2700	ATTAGAGCCGCTGATGCTACCTATATCTGAAATGGAAACTTCACCTCTCTCCACACACTAGAG	2759
Qy	2242	CGGGACCTGCTGTTGAAGGCGACTGCTGCGGTACAGCGGCACGGGCACAGCGCTGGAG	2301
Db	2760	CAAGCACTCACCTACAAGTACTGCTTAAGGTACATGGTTCCTCGGCTCGCTGGAA	2819
Qy	2302	AGCCTGCAGGCTTCCCGGCCATTCCTGGAGCGCTGACGGTGGAGGTCTCTCCGTGGGG	2361
Db	2820	AGAATCCGCAGCTTTAGTCCACTCAAGAAGACCCCTTAACCATCCAGGTTCTTATGCTAGGC	2879
Qy	2362	AAGATGACACGGCCCGGGTCCGCTACTCCTTCTATCTGCCCAAGAGCCTCGGAGGAC	2421
Db	2880	CATGCTCTCCGACCCCAAAATTAATTCACCTACTTTATGAGAGAACAGAGATCATTC	2939
Qy	2422	AAGTCCTCTC 2431	
Db	2940	AACGCCATTC 2949	
RESULT 15			
AAC90067			
ID	AAC90067	standard; DNA; 4180 BP.	
XX			
AC	AAC90067;		
XX			
DT	19-MAR-2001	(first entry)	
XX			
DE	D67076	cdNA clone.	
XX			
KW	METH; metalloproteinase; thrombospondin; angiogenesis inhibition;		
KW	cancer therapy; benign tumour; ocular angiogenic disease;		
KW	rheumatoid arthritis; psoriasis; wound healing; endometriosis;		
KW	vasculogenesis; granulation; hypertrophic scar; nonunion fracture;		
KW	scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;		
KW	coronary collateral; cerebral collateral; arteriovenous malformation;		
KW	ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;		
KW	plaque neovascularisation; telangiectasia; haemophilic joint; EST;		
KW	angiofibroma; fibromuscular dysplasia; expressed sequence tag;		
KW	Crohn's disease; atherosclerosis; birth control; 85.		
XX			
OS	Unidentified.		
XX			
PN	WO200071577-A1.		
XX			
PD	30-NOV-2000.		
XX			
PF	25-MAY-2000; 2000WO-US14462.		
XX			
PR	25-MAY-1999; 99US-0318208.		
PR	20-JUL-1999; 99US-0144882.		
PR	10-AUG-1999; 99US-0147823.		
PR	13-AUG-1999; 99US-0373658.		
PR	22-DEC-1999; 99US-0171503.		
PR	22-FEB-2000; 2000US-0183792.		
XX			
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	(SMK) SMITHKLINE BEECHAM CORP.		
PA	(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.		
PA	(IREU/) IREULA-ARISPE L.		
PA	(HAST/) HASTINGS G A.		
PA	(RUBE/) RUBEN S M.		
PA	(JONA/) JONAK Z L.		
PA	(TRUL/) TRULLI S H.		

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 16:17:29 ; Search time 7234 Seconds
(without alignments)
11477.786 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtctctgtggcctact.....gcgtctctgagccgtgtctga 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rtd.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2851.4	99.9	2853	9	HS315733 Homo sapi
2	2817.6	98.8	2930	6	AX342635 Sequence
3	2817.6	98.8	2937	6	AX319860 Sequence
4	960.4	33.7	157963	9	AP002986 Homo sapi
5	960.4	33.7	172905	2	AC025130 Homo sapi
c	896.2	31.4	170682	2	AC023429 Homo sapi
c	783	27.4	182656	2	AC101990 Mus muscu
8	779	27.3	178764	2	AC126507 Rattus no
9	777.2	27.2	28000	9	AP003459 Homo sapi
10	773.6	27.1	3927	10	BC009667 Mus muscu
11	664.4	23.3	2670	6	E55282 Novel metal
12	664.4	23.3	2670	6	E58655 Novel metal
13	664.4	23.3	3711	9	AF060153 Homo sapi
14	634.4	22.2	4180	10	D67076 Mouse mRNa
15	629.8	22.1	4659	9	AF207664 Homo sapi
16	628.2	22.0	3430	9	AF060152 Homo sapi
17	628.2	22.0	4014	6	E29406 Novel integ
18	628.2	22.0	4447	9	AF170084 Homo sapi
19	627.2	22.0	2184	6	E29966 Human ADAMT
20	626.6	22.0	2853	6	E58656 Novel metal
21	626.6	22.0	4670	9	BC036515 Homo sapi
22	625	21.9	4309	9	AB037767 Homo sapi
23	612.8	21.5	4878	10	AF149118 Rattus no
24	606.4	21.3	2857	10	AF304446 Rattus no
c	597.4	20.9	178764	2	AC126507 Rattus no
26	592	20.8	3638	10	AF175282 Mus muscu
27	583.2	20.4	4193	9	AF148213 Homo sapi
28	583.2	20.4	4301	9	AB014588 Homo sapi
29	581.6	20.4	4407	6	AX376284 Sequence
30	562.6	19.7	3321	10	BC027773 Mus muscu
31	549	19.2	3002	10	AF140673 Mus muscu
32	549	19.2	5533	9	AF142099 Homo sapi
33	417.6	14.6	5139	9	AB037733 Homo sapi
34	417.6	14.6	5808	6	AX319857 Sequence
35	416	14.6	3674	9	AF261918 Homo sapi
36	402.4	14.1	1251	10	AB042271 Rattus no
37	382.4	13.4	2838	9	HS0802617 Homo sapi
38	375.6	13.2	5724	6	AX392458 Sequence
39	375.6	13.2	6060	6	AX392460 Sequence
40	374	13.1	172905	2	AC025130 Homo sapi
41	367.8	12.9	1560	9	AF141293 Homo sapi
42	322	11.3	2554	9	AK094057 Homo sapi
43	319.6	11.2	1365	10	AB042272 Rattus no
44	316	11.1	3308	6	AX480930 Sequence
45	316	11.1	4518	6	AX319858 Sequence

ALIGNMENTS

RESULT 1	HS315733	2853 bp	mRNA	linear	PRI 01-MAR-2002
LOCUS	Homo sapiens mRNA for metalloprotease disintegrin 15 (ADAMTS15 gene)				
DEFINITION	ADAMTS15 gene; disintegrin; metalloprotease; thrombospondin.				
ACCESSION	AJ315733				
VERSION	AJ315733.1	GI:19171175			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Cal.S., Obaya,A.J., Llanazares,M., Garabaya,C., Quesada,V. and Lopez-Otin,C.				

TITLE Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains
JOURNAL Gene 283 (1-2), 49-62 (2002)
MEDLINE 21856482
PUBMED 11867212
REFERENCE 2 (bases 1 to 2853)
AUTHORS Cal.S.
TITLE Direct Submission
JOURNAL Submitted (26-Jun-2001) Cal S., Biochemistry and Molecular Biology, University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN
LOCATION/Qualifiers
FEATURES
source
1. .2853
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. .2853
/gene="ADAMTS15"
CDS
1. .2853
/gene="ADAMTS15"
/codon_start=1
/product="metalloprotease disintegrin 15 with thrombospondin domains"
/protein_id="CAC86014.1"
/db_xref="GI:19171176"
/translation="MLLLGLTLAFAGRTAGSEPEREYVVPVIRLDPDINGRRYYWRG
PEDSGQGLIFQTAFOEDFYHLHTPDALAPFSTEHLGVPLQGLTGGSDLRRCF
YSGDVNAEPDSFAVSLCGFLRGAFYRGAEVVISPLPNASAPAAORNSOGAHLQRR
GVPGSGDPTSGVAGSNPALLRALDPYPRRAGFGESESRERSRKRFVSIPIR
YVETLVADSESVKFGADLEHYLLTLATAARLYRHPSILNPINIVVVLRLDRD
SGPKVTNALTLRNECAWOKLNKRVSDPEWFDAILFTFQDLGCAATCTILGMAD
VTGMDPKRSVIEDDGLPSAFTTAHGLHFNPHDNVYKVEEVEFGYKLRNHHMSP
TLQIDRANPWSACSAIITDFLSDGLDLDQSPKISLPDLPGASYTLSSQCEL
AFVGSKPCPMYQCTKWLCTKAKGMVQOTRHFPWADGTSCGEGKCLKGACVERH
NLNKRVDGSKWAKDYPGCSRTSGGVOLARROCTNPTPANGVKYCEGVRYKRSN
LEPCPSASGKSPREOCFAFNGYNHSTNRLTLAVAWPEKYSVSPRDKCLILRANG
TGFPYVLAKPVVDGTLCSPDSTSVCGQKCIKAGCDNGLSKRPDKCVCGGNKSC
KRYTGLTFPMHGNFVVALPAGASSIDIRQYKGLIGDNLALKNQKGLLNGH
FVYSAYERDLVKSGLLRYSGTAVESLOASRPILFLEPTVEVLSVGRMTPRVRYSF
YLPKREDKSSHPKDPGPGSVLHNSVLSNOVEQDPRPAPRVAGSWGSPCASCG
SGLQKRAVCGSGAGCTVPACDAHRPVEVETACGPEPCTWELSAWSPSKSCGRGFQ
RRSLKCVGHGRLRLDQCNLHRKPOELFCVLRPC"
BASE COUNT 522 a 921 c 903 g 507 t
ORIGIN
Query Match 99.9%; Score 2851.4; DB 9; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTTCTGCTGGGCATCCTAACCTTGGCTTCGCCGGCGCAACCGCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGCATCCTAACCTTGGCTTCGCCGGCGCAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTCCATCCGACTGGACCGCGGACATTAACGGCGCGCTAC 120
DB 61 CCAGAGCGGAGGTAGTCTCCATCCGACTGGACCGCGGACATTAACGGCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCAATT 180
DB 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCAATT 180
QY 181 CAGGAGGACTTTTACCTACACTACCGCGGATCCTCAGTTCTTGGCTCCCGCTTCTCC 240
DB 181 CAGGAGGACTTTTACCTACACTACCGCGGATCCTCAGTTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTACCGGGGCTCTCTTCAGACCTGCGACGC 300
DB 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTACCGGGGCTCTCTTCAGACCTGCGACGC 300
QY 301 TCGTTCTATTCTGGGAGCTGAACCGGAGCCGACTCGTTCGCTGCTGTGAGCCTGTGC 360
DB 301 TCGTTCTATTCTGGGAGCTGAACCGGAGCCGACTCGTTCGCTGCTGTGAGCCTGTGC 360
QY 361 GGGGGGCTCCGGAGGCTTTGGGTACCGAGGCGCGGAGTATGTATAGCCCGCTGCC 420

DB 361 GGGGGGCTCCGGAGGCTTTGGTACCGAGCGCGGAGTAGTATGTATAGCCCGCTGCC 420
QY 421 RATGTAGCGCGCGCGCGCGAGCAACAGCAGGCGCGCACACTTCTCTCCAGCGCGG 480
DB 421 AATGTAGCGCGCGCGCGCGAGCAACAGCAGGCGCGCACACTTCTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGCGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGGCG 540
DB 481 GGTGTTCCGGCGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGGCG 540
QY 541 AACCCCGCATCTTACGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGGAG 600
DB 541 AACCCCGCATCTTACGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCAAGGCTTTCTGTCTATCCCGCGGTACGTG 660
DB 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCAAGGCTTTCTGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGGTGTGTCGCGGACAGTCAATGGTCAAGTTCCACGGCGGGGACTTGAACAT 720
DB 661 GAGACGCTGGTGTGTCGCGGACAGTCAATGGTCAAGTTCCACGGCGGGGACTTGAACAT 720
QY 721 TATCTGTGACGCTGCTGGCAACGCGCGGCGGACTCTACCGCATCCAGCATCCTCAAC 780
DB 721 TATCTGTGACGCTGCTGGCAACGCGCGGCGGACTCTACCGCATCCAGCATCCTCAAC 780
QY 781 CCATCAACATCGTTGTGTGCTCAAGGTGCTGTCTTAGAGATCGTGACTCCGGGCCAAG 840
DB 781 CCATCAACATCGTTGTGTGCTCAAGGTGCTGTCTTAGAGATCGTGACTCCGGGCCAAG 840
QY 841 GTACCGGCAATGCGGCGCTGACGCTGGCAACCTTCTGTGCTGGCGAAGAAAGCTGAAC 900
DB 841 GTACCGGCAATGCGGCGCTGACGCTGGCAACCTTCTGTGCTGGCGAAGAAAGCTGAAC 900
QY 901 AAGTGTGAGCAACACACCCCGAGTACTGGGACACTGCATCTCTTCCACAGCAGGAC 960
DB 901 AAGTGTGAGCAACACACCCCGAGTACTGGGACACTGCATCTCTTCCACAGCAGGAC 960
QY 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGCATGGCTGTGGTGGTACCATGTGTGAC 1020
DB 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGCATGGCTGTGGTGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGTCTATTGAGGACGATGGCTTCCATCAGCTTCCACCATGCC 1080
DB 1021 CCCAAGAGAGCTGCTCTGTCTATTGAGGACGATGGCTTCCATCAGCTTCCACCATGCC 1080
QY 1081 CACGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGAGTG 1140
DB 1081 CACGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGAGAGCTCCGAGCAACACATGATGCCCGACCTCATCCAGATCGACCGTGCC 1200
DB 1141 TTTGGAGAGCTCCGAGCAACACATGATGCCCGACCTCATCCAGATCGACCGTGCC 1200
QY 1201 AACCCCTGTGAGCTGCGAGTCTGCATCATCACGACTTCTTGACAGCGGCGACGGT 1260
DB 1201 AACCCCTGTGAGCTGCGAGTCTGCATCATCACGACTTCTTGACAGCGGCGACGGT 1260
QY 1261 GACTGCTCTTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
DB 1261 GACTGCTCTTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCTGAGCCAGTGGCGGCTTTTGGGCTGGGCTTCCAAAGCCCTGTCT 1380
DB 1321 AGCTACACCTGAGCCAGTGGCGGCTTTTGGGCTGGGCTTCCAAAGCCCTGTCT 1380
QY 1381 TACATCAGTACTGCAACAGCTGTGTGTCACCGGGAAGCCCAAGGACAGATGTTGTC 1440
DB 1381 TACATCAGTACTGCAACAGCTGTGTGTCACCGGGAAGCCCAAGGACAGATGTTGTC 1440
QY 1441 CAGACCGGCACTTCCCTTGGGCGGATGACACAGCTGTGGCGGAGGCAAGCTTGCCTC 1500
DB 1441 CAGACCGGCACTTCCCTTGGGCGGATGACACAGCTGTGGCGGAGGCAAGCTTGCCTC 1500

Db 1441 CAGACCCGCCACTTCCCTGGGCCGATGGCCACCAAGCTGTGGCGAGGGCAAGCTCTGCCTC 1500
 QY 1501 AAGGGGCTCGTGGAGAGACAACTCAACAGCACAGGGTGGATGTCTCTGGGC 1560
 Db 1501 AAGGGGCTCGTGGAGAGACAACTCAACAGCACAGGGTGGATGTCTCTGGGC 1560
 QY 1561 AATGGGATCCCTATGGCCCTCTGTCGCGCACATGTGGTGGGGCGCTGCAAGTGGCCAGG 1620
 Db 1561 AATGGGATCCCTATGGCCCTCTGTCGCGCACATGTGGTGGGGCGCTGCAAGTGGCCAGG 1620
 QY 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTGAAGGTG 1680
 Db 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTGAAGGTG 1680
 QY 1681 AATACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCTCCGGAAGAGTTCGGG 1740
 Db 1681 AATACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCTCCGGAAGAGTTCGGG 1740
 QY 1741 GAGGAGCTGTGAGGCTTCAACGGCTACACACAGCACCAACCGGCTCAGCTCTCGCC 1800
 Db 1741 GAGGAGCTGTGAGGCTTCAACGGCTACACACAGCACCAACCGGCTCAGCTCTCGCC 1800
 QY 1801 GTGGATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTCAAGCTCATCTGC 1860
 Db 1801 GTGGATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTCAAGCTCATCTGC 1860
 QY 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGCCACCAAGTGTGGAGGGCACGCTG 1920
 Db 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGCCACCAAGTGTGGAGGGCACGCTG 1920
 QY 1921 TGCTCTCTGACTCCACCTCCCTGTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGAT 1980
 Db 1921 TGCTCTCTGACTCCACCTCCCTGTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGAT 1980
 QY 1981 GGGAACTGGGTCCNAGNAGAGATTCAGCAAGTGTGGGTGTGTGGGGAGACAATAAG 2040
 Db 1981 GGGAACTGGGTCCNAGNAGAGATTCAGCAAGTGTGGGTGTGTGGGGAGACAATAAG 2040
 QY 2041 AGTGTCAAGAAGTGTGACTGGCTTCTACCAAGCCCATGATGGCTACAAATTCGTGGTG 2100
 Db 2041 AGTGTCAAGAAGTGTGACTGGCTTCTACCAAGCCCATGATGGCTACAAATTCGTGGTG 2100
 QY 2101 GCATCCCGGAGGCGCTCAAGCATCGACATCCGCGCAGCGGGTTCACAAAGGCTGATC 2160
 Db 2101 GCATCCCGGAGGCGCTCAAGCATCGACATCCGCGCAGCGGGTTCACAAAGGCTGATC 2160
 QY 2161 GGGGATGCAACTTACCTGGCTCTGAGAACAGCCAAAGTACCTGCTCAACGGGCAT 2220
 Db 2161 GGGGATGCAACTTACCTGGCTCTGAGAACAGCCAAAGTACCTGCTCAACGGGCAT 2220
 QY 2221 TTGCTGGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGGGCAGTCTGTCGCGTACAGC 2280
 Db 2221 TTGCTGGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGGGCAGTCTGTCGCGTACAGC 2280
 QY 2281 GGACGGGACACGGGTGGAGAGCTGCGAGGCTTCCCGGCCCATCTCCGCGCTGACC 2340
 Db 2281 GGACGGGACACGGGTGGAGAGCTGCGAGGCTTCCCGGCCCATCTCCGCGCTGACC 2340
 QY 2341 GTGGAGTCTCTCTCGGTGGGAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
 Db 2341 GTGGAGTCTCTCTCGGTGGGAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
 QY 2401 CCCAAGAGCTCGGAGAGACAACTCTCTATCCCAAGGACCCCGGGACCCCTCTGTGTC 2460
 Db 2401 CCCAAGAGCTCGGAGAGACAACTCTCTATCCCAAGGACCCCGGGACCCCTCTGTGTC 2460
 QY 2461 TTGCACACAGGCTCTCAGCTCTCCAAACAGGTGGAGCGCGGACGACAGCCGCCCT 2520
 Db 2461 TTGCACACAGGCTCTCAGCTCTCCAAACAGGTGGAGCGCGGACGACAGCCGCCCT 2520
 QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGGTGCTCCCGGAGGTGGCGAGTGGGCTGCAG 2580
 Db 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGGTGCTCCCGGAGGTGGCGAGTGGGCTGCAG 2580

QY 2581 AAGCGGCGGTGACTGTCCGGCTCCGCGGCGAGCGACGGTCCCTGCTGTGA 2640
 Db 2581 AAGCGGCGGTGACTGTCCGGGCTCCGCGGCGAGCGACGGTCCCTGCTGTGA 2640
 QY 2641 GCCCATCGGCGGTGGAGACACAAAGCTTGGGGAGCCCTGCCCACTGGGAGCT 2700
 Db 2641 GCCCATCGGCGGTGGAGACACAAAGCTTGGGGAGCCCTGCCCACTGGGAGCT 2700
 QY 2701 GCTGTGCTCACCCTGTCTCAAGAGCTCGCGCGGGATTTCAGAGGCGCTCACTCAATGT 2760
 Db 2701 GCTGTGCTCACCCTGTCTCAAGAGCTCGCGCGGGATTTCAGAGGCGCTCACTCAATGT 2760
 QY 2761 GTGGCGCAGGAGCGGCTGTGTGGCGCGGACGAGTGCATTCGACCGAGCCGAG 2820
 Db 2761 GTGGCGCAGGAGCGGCTGTGTGGCGCGGACGAGTGCATTCGACCGAGCCGAG 2820
 QY 2821 GAGCTGGACTTCTCGCTCCTGAGGCGGTGCTGA 2853
 Db 2821 GAGCTGGACTTCTCGCTCCTGAGGCGGTGCTGA 2853

RESULT 2
 AX342635
 LOCUS AX342635 2930 bp DNA linear Pat 12-JAN-2002
 DEFINITION Sequence 32 from Patent WO0198468.
 ACCESSION AX342635
 VERSION AX342635.1 GI:18152032
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Yue.H., Elliott.V.S., Gandhi.A.R., Lal.P., Au-Young.J.,
 Tribouley.C.M., Deleage.A.M., Baughn.M.R., Nguyen.D.P., Lee.E.A.,
 Hatala.A., Khan.F.A., Walla.N.K., Yao.M.G., Lu.D.A., Patterson.C.,
 Tang.Y.T., Walsh.R.T., Azimzai.Y., Ramkumar.J., Xu.Y. and Reddy.R.
 Patent: WO 0198468-A 32 27-DEC-2001;
 Incyte Genomics, Inc. (US)

FEATURES
 Location/Qualifiers
 source 1..2930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7473089CBI"

BASE COUNT 529 a 958 c 931 g 512 t

ORIGIN

Query Match 98.8%; Score 2817.6; DB 6; Length 2930;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2846; Conservative 0; Mismatches 4; Indels 6; Gaps 2;

QY 1 ATGCTTCTCTGGGATCTTAACCTTGGCTTTCGCGGGCGAAGCGTGGAGGCTCTGAG 60
 Db 75 ATGCTTCTCTGGGATCTTAACCTTGGCTTTCGCGGGCGAAGCGTGGAGGCTCTGAG 134
 QY 61 CCAGAGCGGAGGTAGTCTTCCCATTCGACTTGGACCGGACATTAAGCGCGCGCTAC 120
 Db 135 CCAGAGCGGAGGTAGTCTTCCCATTCGACTTGGACCGGACATTAAGCGCGCGCTAC 194
 QY 121 TACTGCGGGGTCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180
 Db 195 TACTGCGGGGTCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 254
 QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCCTTGGCTCCCGCTTCC 240
 Db 255 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCCTTGGCTCCCGCTTCC 314
 QY 241 ACTGAGCATCTGGGCTGCCCTCCAGGGCTCAGCGGGGCTCTTCAGACTGCGAGCGC 300
 Db 315 ACTGAGCATCTGGGCTGCCCTCCAGGGCTCAGCGGGGCTCTTCAGACTGCGAGCGC 374
 QY 301 TGCTTCTATTCTGGGACGCTGAACCGGAGCGGACTGCTTCTGCTGCTGTGAGGCTGTGC 360


```
QY 2515 CCCCTGCACGCTGGTGGCTGGAGCTGGGGGCGCTGCTCCGGAGCTCGGGCAGTGGC 2574
Db 2595 CCCCTGCACGCTGGTGGCTGGAGCTGGGGGCGCTGCTCCGGAGCTCGGGCAGTGGC 2654
QY 2575 CTGCGAGAGCGGGCGGTGGAGCTGGGGGCTCCGCGGGGAGCGCAGCGTCCCTGGCTGT 2634
Db 2655 CTGCGAGAGCGGGCGGTGGAGCTGGGGGCTCCGCGGGGAGCGCAGCGTCCCTGGCTGT 2714
QY 2635 GATGAGCGCCATCGCCCGCTGGAGACACAAGCCTCGGGGAGCGCTCCGCCACCTGGGAG 2694
Db 2715 GATGAGCGCCATCGCCCGCTGGAGACACAAGCCTCGGGGAGCGCTCCGCCACCTGGGAG 2774
QY 2695 CTCAGCGCTGTGTACACCTGTCTCAAGAGCTGGCGGCGGGGATTTACAGAGCGCTCACTC 2754
Db 2775 CTCAGCGCTGTGTACACCTGTCTCAAGAGCTGGCGGCGGGGATTTACAGAGCGCTCACTC 2834
QY 2755 AAGTGTGGGCGACGGAGCGGCTGTGCGCGGGAGCAGTGGCGGACCACTTGACCGCAAG 2814
Db 2835 AAGTGTGGGCGACGGAGCGGCTGTGCGCGGGAGCAGTGGCGGACCACTTGACCGCAAG 2894
QY 2815 CCCAGGAGCTGGACTTCTGGCTCTGAGGCGGTGC 2850
Db 2895 CCCAGGAGCTGGACTTCTGGCTCTGAGGCGGTGC 2930

RESULT 3
AX319860
LOCUS AX319860 2937 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 24 from Patent WO0183782.
ACCESSION AX319860
VERSION AX319860.1 GI:17901450
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
AUTHORS Plowman,G.D., Whyte,D., Sudarsanam,S., Manning,S., Caenepeel,S. and
Payne,V.
TITLE Novel proteases
JOURNAL Patent: WO 0183782-A 24 08-NOV-2001;
Sugen, Inc. (US)
FEATURES
source
Location/Qualifiers
1..2937
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 531 a 958 c 936 g 512 t
ORIGIN

Query Match 98.8%; Score 2817.6; DB 6; Length 2937;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2849; Conservative 0; Mismatches 4; Indels 9; Gaps 2;

QY 1 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCGGGCAACCGCTGGAGGCTCTGAG 60
Db 76 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCGGGCAACCGCTGGAGGCTCTGAG 135
QY 61 CCAGAGCGGAGGTAGTGTCTCCATCCGACTGGACCCGGACATTAACGGCCGCGGCTAC 120
Db 136 CCAGAGCGGAGGTAGTGTCTCCATCCGACTGGACCCGGACATTAACGGCCGCGGCTAC 195
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCAGAGCATTT 180
Db 196 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCAGAGCATTT 255
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCCTGGGCTCCCGCTTCTCC 240
Db 256 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCCTGGGCTCCCGCTTCTCC 315
QY 241 ACTGAGCATCTGGGGTCCCTCCAGGGGCTCACCAGGGGCTCTTCAGACCTCGGAGCG 300
Db 316 ACTGAGCATCTGGGGTCCCTCCAGGGGCTCACCAGGGGCTCTTCAGACCTCGGAGCG 375
```

```
QY 301 TCGTTCTATTCTGGGAGCTGAAGCCGAGCGGACTCGTTTCGCTGCTGTGAGCTGTGC 360
Db 376 TCGTTCTATTCTGGGAGCTGAAGCCGAGCGGACTCGTTTCGCTGCTGTGAGCTGTGC 435
QY 361 GGGGGGCTCCGCGGAGCTTTGGCTACCGAGGCGCGAGTATGATCATTAGCCGCTGCC 420
Db 436 GGGGGGCTCCGCGGAGCTTTGGCTACCGAGGCGCGAGTATGATCATTAGCCGCTGCC 495
QY 421 AATGCTAGCGCGCGCGGCGCAGCGCAACAGCAGGCGGCACACCTTCTCCAGCGCGG 480
Db 496 AATGCTAGCGCGCGCGGCGCAGCGCAACAGCAGGCGGCACACCTTCTCCAGCGCGG 555
QY 481 GGTGTTCGCGGGGCGGCTTCGCGAGACCCACCTCTCTCGCTCGGGGTGGCCCTCGGCTG 540
Db 556 GGTGTTCGCGGGGCGGCTTCGCGAGACCCACCTCTCTCGCTCGGGGTGGCCCTCGGCTG 615
QY 541 AACCCCGCATCTACGGGCGCTTGACACCTTACAGCGCGCGGCGGGGCTTCGGGAG 600
Db 616 AACCCCGCATCTACGGGCGCTTGACACCTTACAGCGCGCGGCGGGGCTTCGGGAG 675
QY 601 AGTCGTAGCGCGCAGGTCTGGGCGCGCCAAAGCGTTTCGTGTCTATCCCGCGGTACGTG 660
Db 676 AGTCGTAGCGCGCAGGTCTGGGCGCGCCAAAGCGTTTCGTGTCTATCCCGCGGTACGTG 735
QY 661 GAGACGCTGGTGGTCGCGGAGAGTCAATGCTCAAGTTCACGCGCGGCGGCTTCGGGAG 720
Db 736 GAGACGCTGGTGGTCGCGGAGAGTCAATGCTCAAGTTCACGCGCGGCGGCTTCGGGAG 795
QY 721 TATCTGTGACGCTCTCTGGCAACGCGCGGCGGACTCTACCGCCATCCCGAGCATCTCAAC 780
Db 796 TATCTGTGACGCTCTCTGGCAACGCGCGGCGGACTCTACCGCCATCCCGAGCATCTCAAC 855
QY 781 CCCATCAACATCGTTGGTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCAAG 840
Db 856 CCCATCAACATCGTTGGTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCAAG 915
QY 841 GTCACCGGCAATGCGGCGCTGACGCTGGCGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 900
Db 916 GTCACCGGCAATGCGGCGCTGACGCTGGCGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 975
QY 901 AAAGTAGTGAACAAGACCCCGAGTACTGGGAGACTGCCATCTCTTACACGAGCAAG 960
Db 976 AAAGTAGTGAACAAGACCCCGAGTACTGGGAGACTGCCATCTCTTACACGAGCAAG 1035
QY 961 CTGTGTGAGGACCAACACCTGTGACACCTCGGCGGCTGATGTGGGTACCATGTGAC 1020
Db 1036 CTGTGTGAGGACCAACACCTGTGACACCTCGGCGGCTGATGTGGGTACCATGTGAC 1095
QY 1021 CCCAAGAGAGCTGCTGTCTATTGAGGAGGATGGGCTTCCATCAGCCTTCCACCAAG 1080
Db 1096 CCCAAGAGAGCTGCTGTCTATTGAGGAGGATGGGCTTCCATCAGCCTTCCACCAAG 1155
QY 1081 CAGGAGCTGGGCGCAGTGTTCACATGCCATGCCATGAAAGTCTGTGAGGAGGTG 1140
Db 1156 CAGGAGCTGGGCGCAGTGTTCACATGCCATGCCATGAAAGTCTGTGAGGAGGTG 1215
QY 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACCTGTC 1200
Db 1216 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACCTGTC 1275
QY 1201 AACCCCTGTGAGCTGCGAGTGTGCCATCATACCGACTTCCCTGGACAGCGGCAAG 1260
Db 1276 AACCCCTGTGAGCTGCGAGTGTGCCATCATACCGACTTCCCTGGACAGCGGCAAG 1335
QY 1261 GACTGCTCTCTGGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGTGCGGGGCG 1320
Db 1336 GACTGCTCTCTGGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGTGCGGGGCG 1395
QY 1321 AGCTACACCTCAGCCAGGAGTGGAGCTGGCTTTTGGGTGGGCTCCAAAGCCCTGTCT 1380
Db 1396 AGCTACACCTCAGCCAGGAGTGGAGCTGGCTTTTGGGTGGGCTCCAAAGCCCTGTCT 1455
```

```
QY 1381 TACATCAGTACTGCACCAAGCTGTGTGCACCGGGAAGCCCAAGGACAGATGTTGTC 1440
|||||
Db 1456 TACATCAGTACTGCACCAAGCTGTGTGCACCGGGAAGCCCAAGGACAGATGTTGTC 1515
|||||
QY 1441 CAGACCCGCCACTTCCCTGGGCCGATGACCAAGCTGTGGCGAGGGCAAGCTCTGCCTC 1500
|||||
Db 1516 CAGACCCGCCACTTCCCTGGGCCGATGACCAAGCTGTGGCGAGGGCAAGCTCTGCCTC 1575
|||||
QY 1501 AAAGGGCCCTGCGTGAGAGACACACCTCAACAACACAGGTCGATGCTTCTGGGCC 1560
|||||
Db 1576 AAAGGGCCCTGCGTGAGAGACACACCTCAACAACACAGGTCGATGCTTCTGGGCC 1635
|||||
QY 1561 AAATGGGATCCCTATGCCCCCTCTCGCCACATGTGTGGGGCGCTGCAGCTGGCCAGG 1620
|||||
Db 1636 AAATGGGATCCCTATGCCCCCTCTCGCCACATGTGTGGGGCGCTGCAGCTGGCCAGG 1695
|||||
QY 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGGCAAGTACTGCAGGAGAGTGAGGGTG 1680
|||||
Db 1696 AGGAGTGCACCAACCCACCCCTGCCAACGGGGGCAAGTACTGCAGGAGAGTGAGGGTG 1755
|||||
QY 1681 AAATACCGATCCCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCCTCCGGAAGAGCTTCCGG 1740
|||||
Db 1756 AAATACCGATCCCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCCTCCGGAAGAGCTTCCGG 1815
|||||
QY 1741 GAGGAGCAGTGTGAGGGCTTTCAACGGCTACACCAACAGACCAACCGGCTCACTCTCGCC 1800
|||||
Db 1816 GAGGAGCAGTGTGAGGGCTTTCAACGGCTACACCAACAGACCAACCGGCTCACTCTCGCC 1875
|||||
QY 1801 GTGCATGGGTGCCCAAGTACTCGGGGTGTCTCCCGGGACAAGTGCAAGCTCACTCTGC 1860
|||||
Db 1876 GTGCATGGGTGCCCAAGTACTCGGGGTGTCTCCCGGGACAAGTGCAAGCTCACTCTGC 1935
|||||
QY 1861 CGAGCCAATGGCACTGTCTACTTCTATGTCTGTCGACCCAA---GGTGTGGACGGCACG 1917
|||||
Db 1936 CGAGCCAATGGCACTGTCTACTTCTATGTCTGTCGACCCAAAGGTGTGGTGGACGGCACG 1995
|||||
QY 1918 CTGTGCTCTCTGACTCCACTCCCTCGTGTGTCTCAAGGCAAGTGCATCAAGGCTGCTGT 1977
|||||
Db 1996 CTGTGCTCTCTGACTCCACTCCCTCGTGTGTCTCAAGGCAAGTGCATCAAGGCTGCTGT 2055
|||||
QY 1978 CATGGAACTGGGCTCCAAAGAGAGATTCGACAAGTGGGTGTGTGGGGGAGACAAT 2037
|||||
Db 2056 GATGGAACTGGGCTCCAAAGAGAGATTCGACAAGTGGGTGTGTGGGGGAGACAAT 2115
|||||
QY 2038 AAGAGCTGCAAGAAGTGTGACTGCTCTTCAACAAGCCCATGATGCTACAAATTCGTG 2097
|||||
Db 2116 AAGAGCTGCAAGAAGTGTGACTGCTCTTCAACAAGCCCATGATGCTACAAATTCGTG 2175
|||||
QY 2098 GTGGCCATCCCGCAGGGGCTCAAGCATCGACATCCGCCAGCGGTTACAAGGGCTG 2157
|||||
Db 2176 GTGGCCATCCCGCAGGGGCTCAAGCATCGACATCCGCCAGCGGTTACAAGGGCTG 2235
|||||
QY 2158 ATCGGGATGACAATCTACTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGG 2217
|||||
Db 2236 ATCGGGATGACAATCTACTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGG 2295
|||||
QY 2218 CATTTCTGTGTGTGCGGGTGGAGCGGACCTGTGTGTGAAGGCACTCTGTCGGGTAC 2277
|||||
Db 2296 CATTTCTGTGTGTGCGGGTGGAGCGGACCTGTGTGTGAAGGCACTCTGTCGGGTAC 2355
|||||
QY 2278 AGCGCACGGGCACAGCGGTGGAGGCTGACAGCTTCCCGGCCCATCTCTGGAGCGGCTG 2337
|||||
Db 2356 AGCGCACGGGCACAGCGGTGGAGGCTGACAGCTTCCCGGCCCATCTCTGGAGCGGCTG 2415
|||||
QY 2338 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2397
|||||
Db 2416 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2475
|||||
QY 2398 CTGCCCAAAGACCTTCGGGAGGACAAGTCTCTCATGCCAAGGACCCCGG-----GGA 2451
|||||
Db 2476 CTGCCCAAAGACCTTCGGGAGGACAAGTCTCTCATGCCCGCACCCCGGGGAGGAGA 2535
|||||
QY 2452 CCCTCTGTCTTGACACACAGCGTCTCTCAGCCCTCTCCACACAGGTGGAGCACCGGACGAC 2511
|||||
```

```
Db 2536 CCCTCTGTCTTGACACACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCACCGGACGAC 2595
|||||
QY 2512 AGGCCCTCTGACAGCTGGGTGGCTGGAGCTGGGGGCCGTCTCCGGAGCTCGGCGAGT 2571
|||||
Db 2596 AGGCCCTCTGACAGCTGGGTGGCTGGAGCTGGGGGCCGTCTCCGGAGCTCGGCGAGT 2655
|||||
QY 2572 GGCCTGCAGAAAGCGGGGGGTGACTGTGGGGGTCTCCGCCGGGAGCGACAGTTCCTTGCC 2631
|||||
Db 2656 GGCCTGCAGAAAGCGGGGGTGGACTGTCCGGGGTCTCCGCCGGGAGCGACAGCTTCCCTGCC 2715
|||||
QY 2632 TGTGATGAGCCCATCGGCCGTGGAGACACAGCCTGGGGGAGCCCTGCCACCTGG 2691
|||||
Db 2716 TGTGATGAGCCCATCGGCCGTGGAGACACAGCCTGGGGGAGCCCTGCCACCTGG 2775
|||||
QY 2692 GAGCTCAGCGCTGTGTACCTGTCTCAAGAGCTGTCCGGCGGGGATTCAGAGCGGCTCA 2751
|||||
Db 2776 GAGCTCAGCGCTGTGTACCTGTCTCAAGAGCTGTCCGGCGGGGATTCAGAGCGGCTCA 2835
|||||
QY 2752 CTCAGTGTGTGGGCCAGGAGCGGCTGTGCCCGGGAGCCAGTGCACATTGCACCGC 2811
|||||
Db 2836 CTCAGTGTGTGGGCCAGGAGCGGCTGTGCCCGGGAGCCAGTGCACATTGCACCGC 2895
|||||
QY 2812 AAGCCCCAGGAGCTGACTTCTGCTGAGGCGGTGCTGA 2853
|||||
Db 2896 AAGCCCCAGGAGCTGACTTCTGCTGAGGCGGTGCTGA 2937
|||||

RESULT 4
AP002986 157963 bp DNA linear p: 08-MAR-2002
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-121M22,
DEFINITION complete sequence.
ACCESSION AP002986
VERSION AP002986.2 GI:19263031
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-121M22.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H.
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seon,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seon,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suhiro-chou, Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gs.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Mar 7, 2002 this sequence version replaced gi:11559301.
FEATURES
source
1..157963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-121M22"

BASE COUNT 39742 a 36722 c 37247 g 44252 t
ORIGIN

Query Match 33.7%; Score 960.4; DB 9; Length 157963;
Best Local Similarity 98.4%; Pred. No. 6.7e-163;
Matches 970; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGCTTCTCTGGGCATCTAACCCTGGCTTTCGCGGGGCGGAGCCGCTGGAGGCTCTGAG 60
|||||
Db 134115 ATGCTTCTCTGGGCATCTAACCCTGGCTTTCGCGGGGCGGAGCCGCTGGAGGCTCTGAG 134174
|||||
```

```
QY 61 CCAGAGCGGAGTAGTCTGTTCCATCCGACTCGAGCCGACATTAACGGCCGCGCTAC 120
|||||
Db 134175 CCAGAGCGGAGTAGTCTGTTCCATCCGACTCGAGCCGACATTAACGGCCGCGCTAC 134234
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCACAGCATTT 180
|||||
Db 134235 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCACAGCATTT 134294
QY 181 CAGGAGGACTTTTACTACACTGACGCGGATGCTCAGTTCTTGTTGGCTCCCGCTTCCTCC 240
|||||
Db 134295 CAGGAGGACTTTTACTACACTGACGCGGATGCTCAGTTCTTGTTGGCTCCCGCTTCCTCC 134354
QY 241 ACTGAGCATCTGGGCTCCCTCCAGGGCTCACGGGGCTCTTCAGACTCGGAGCG 300
|||||
Db 134355 ACTGAGCATCTGGGCTCCCTCCAGGGCTCACGGGGCTCTTCAGACTCGGAGCG 134414
QY 301 TCGTTCTATTCTGGGAGGTGAACCCGAGCGGACTCGTTCGCTGCTGTCGAGCTGTGC 360
|||||
Db 134415 TCGTTCTATTCTGGGAGGTGAACCCGAGCGGACTCGTTCGCTGCTGTCGAGCTGTGC 134474
QY 361 GGGGGCTCCCGGAGCCTTTGGCTTACGAGGCGCGGAGTATGTCATTTAGCCCGCTGCGCC 420
|||||
Db 134475 GGGGGCTCCCGGAGCCTTTGGCTTACGAGGCGCGGAGTATGTCATTTAGCCCGCTGCGCC 134534
QY 421 AATGCTAGCGCGCGCGCGGAGCAACAGCGGCGGAGTATGTCATTTAGCCCGCTGCGCC 480
|||||
Db 134535 AATGCTAGCGCGCGCGGAGCAACAGCGGCGGAGTATGTCATTTAGCCCGCTGCGCC 134594
QY 481 GGTGTTCCGGGGCGGCTTCCGGAGACCCCACTCTCGCTCGGGGGTGGCTCGGGGTGG 540
|||||
Db 134595 GGTGTTCCGGGGCGGCTTCCGGAGACCCCACTCTCGCTCGGGGGTGGCTCGGGGTGG 134654
QY 541 AACCCCGCCATCTACGGGCGCTTGAGCCCTTACAGCGCGCGGGCGGCTTCGGGGAG 600
|||||
Db 134655 AACCCCGCCATCTACGGGCGCTTGAGCCCTTACAGCGCGCGGGCGGCTTCGGGGAG 134714
QY 601 AGTCGTAGCGCGCGGAGTCTGGGCGGCCCAAGGTTTCGTTGCTATCCGCGGTAGTG 660
|||||
Db 134715 AGTCGTAGCGCGCGGAGTCTGGGCGGCCCAAGGTTTCGTTGCTATCCGCGGTAGTG 134774
QY 661 GAGAGCTGGTGTGGCGGACAGTCAATGGTCAAGTTTCACGCGCGGCGGACTCGGAACAT 720
|||||
Db 134775 GAGAGCTGGTGTGGCGGACAGTCAATGGTCAAGTTTCACGCGCGGCGGACTCGGAACAT 134834
QY 721 TATCTGCTGAGCTGCTGGCAACGCGCGGAGTCTACCGCCATCCAGCATCCTCAAC 780
|||||
Db 134835 TATCTGCTGAGCTGCTGGCAACGCGCGGAGTCTACCGCCATCCAGCATCCTCAAC 134894
QY 781 CCCATCAACATCTGTTGGTCAAGTGTGCTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
|||||
Db 134895 CCCATCAACATCTGTTGGTCAAGTGTGCTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 134954
QY 841 GTCACCGGCAATCGGGCCCTGACGCTCGCAACTTCTGTCCTGGCAAGAGAGCTGAAC 900
|||||
Db 134955 GTCACCGGCAATCGGGCCCTGACGCTCGCAACTTCTGTCCTGGCAAGAGAGCTGAAC 135014
QY 901 AAAGTACTGACAGCACCCCGAGTACTGGACACTGCCATCCTCTTCAACAGGAGGAC 960
|||||
Db 135015 AAAGTACTGACAGCACCCCGAGTACTGGACACTGCCATCCTCTTCAACAGGAGGAC 135074
QY 961 CTGTGTGGAGCCACCATCTGTGACAC 986
|||||
Db 135075 AGTTGATCGCGTCACTTTGCACCC 135100
```

```
RESULT 5
AC025130/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-211H6 map 11, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
AC025130
VERSION AC025130.2 GI:7387384
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
```

SOURCE
ORGANISM

human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 172905)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-211H6

Unpublished

2 (bases 1 to 172905)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,

Boguslavsky,L., Boukhgaltier,B., Brown,A., Burkett,G.,

Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Kaitas,A.,

Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,

Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,

McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,

Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 172905)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,

Boguslavsky,L., Boukhgaltier,B., Brown,A., Burkett,G.,

Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,

Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,

Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,

McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,

Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 3, 2000 this sequence version replaced gi:7158941.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7808

Center clone name: 211_H6

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
 Consensus quality: 161899 bases at least Q40.
 Consensus quality: 167402 bases at least Q30.
 Consensus quality: 165801 bases at least Q20.
 Insert size: 177000; agarose-fp
 Insert size: 171405; sum-of-contigs
 Quality coverage: 4.1 in Q20 bases; agarose-fp
 Quality coverage: 4.2 in Q20 bases; sum-of-con-

[illegible]


```

QY 721 TATCTGCTGACGCTGTGGCAACGGCGGAGCTTACCGCATCCAGCATCCTCAAC 780
|||||
Db 78863 TATCTGCTGACGCTGTGGCAACGGCGGAGCTTACCGCATCCAGCATCCTCAAC 78804
QY 781 CCATCAACATGTTGTGGTCAAGTGCTGCTTCTTAGAGATCGTGACTCGGGGCCAAG 840
|||||
Db 78803 CCATCAACATGTTGTGGTCAAGTGCTGCTTCTTAGAGATCGTGACTCGGGGCCAAG 78744
QY 841 GTCACGGCAATCGGCCCTGACGCTGGCAACTTCTGCTGGCAGAGAACTGAAC 900
|||||
Db 78743 GTCACGGCAATCGGCCCTGACGCTGGCAACTTCTGCTGGCAGAGAACTGAAC 78684
QY 901 AAGTGTAGTGAACACCCCGAGTACTGGGACACTGCGATCTCTTCCACGAGCAGAC 960
|||||
Db 78683 AAGTGTAGTGAACACCCCGAGTACTGGGACACTGCGATCTCTTCCACGAGCAGAC 78624
QY 961 CTGTGTGAGGACCAACACTGTGACAC 986
|||||
Db 78623 AGTTGATCTGCCGTCACTTTGCACCC 78598

RESULT 6
AC023429 170682 bp DNA linear HTG 10-APR-2001
LOCUS Homo sapiens chromosome 11 clone RP11-121M22, WORKING DRAFT
DEFINITION
AC023429
AC023429 GI:13569974
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ABOLA, A.P., Bruno D., Conn, L., Dela Rosa, M., Faulkner, D.,
Fiederspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,
Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,
Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D.,
Southwick, A.M., Webb, C., Wilhelmy, J., Yu, S. and Davis, R.W.
Unpublished
2 (bases 1 to 170682)
Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federspiel, N.,
Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J.,
Yu, S. and Davis, R.W.
Direct Submission
Submitted (14-FEB-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Apr 10, 2001 this sequence version replaced gi:13562078.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center
Center code: SDSTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: 837
Center clone name: RP11-121M22
----- Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye-primer; 12% of reads
Assembly: Dye-terminator Big Dye; 86% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165770 bases at least Q40
Consensus quality: 167255 bases at least Q30
Consensus quality: 167918 bases at least Q20
Insert size: 172423; agarose-fp
Insert size: 170082; sum-of-ctnigs
Quality coverage: 9.0x in Q20 bases; agarose-fp
Quality coverage: 9.1x in Q20 bases; sum-of-ctnigs.
* NOTE: This is a 'working draft' sequence. It currently

```

```

* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1856: contig of 1856 bp in length
* 1857 1956: gap of unknown length
* 1957 7065: contig of 5109 bp in length
* 7066 7165: gap of unknown length
* 7166 21089: contig of 13924 bp in length
* 21090 21189: gap of unknown length
* 21190 34145: contig of 12956 bp in length
* 34146 48562: gap of unknown length
* 48563 48662: contig of 14317 bp in length
* 48663 109179: gap of unknown length
* 109180 109279: contig of 60517 bp in length
* 109280 170682: gap of unknown length
* 170682 61403 bp in length.
FEATURES
Location/Qualifiers
Source
1..170682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-121M22"
/clone_lib="RPC1 human BAC library 11"
misc_feature
1..1856
/note="assembly_name:Contig25"
1957..7065
/note="assembly_name:Contig26"
7166..21089
/note="assembly_name:Contig27"
21190..34145
/note="assembly_name:Contig28"
34246..48562
/note="assembly_name:Contig29"
clone_end:SP6
48663..109179
/note="assembly_name:Contig30"
109280..170682
/note="assembly_name:Contig31"
clone_end:77"
BASE COUNT 43947 a 39732 c 39514 g 46884 t 605 others
ORIGIN
Query Match 31.4%; Score 896.2; DB 2; Length 170682;
Best Local Similarity 96.2%; Pred. NO. 2.2e-151;
Matches 950; Conservative 0; Mismatches 34; Indels 4; Gaps 3;
QY 1 ATGCTTCTCTGGGATCTCTAACCCTGGCTTTTCGGCGGCGAACCCTGGAGGCTCTGAG 60
|||||
Db 98460 ATGCTTCTCTGGGATCTCTAACCCTGGCTTTTCGGCGGCGAACCCTGGAGGCTCTGAG 98519
QY 61 CCAG-AGCGGGAGGTA-GTCGTTCCCATCCGACTGGACCCGGACATTAAACGCCGCCGCT 118
|||||
Db 98520 CCAGCAGCGGAGGAGTACGTCGTTCCCATTCGACTGGACCCGGACATTAAACGCTTTCCT 98579
QY 119 ACTACTGGCGGGGTCCCGAGGAGTCCGGGGATCAGGGACTCATTTTTCAGATCAGCAT 178
|||||
Db 98580 TTATCTGGCGGGTACC--GAGGACCCCGGGGTACAGNACTCATTTTTCAGATCAGCAT 98637
QY 179 TTCAGGAGGACTTTTACCTACACCTGACCGCGATGCTCAGTTCTTGCTCCCGCTTCT 238
|||||
Db 98638 TTCAGGAGGACTTTTACCTACACCTGACCGCGATGCTCAGTTCTTGCTCCCGCTTCT 98697
QY 239 CCACTGAGCATCTGGCGGCTCCCGCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTTGCAC 298
Db 98698 CCACTGAGCATCTGGCGGCTCCCGCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTTGCAC 98757
QY 299 GCTGCTTCTTCTGGGAGCGTGAACGCCGAGCGGAGTCTGCTGCTGAGACCTGT 358
|||||
Db 98758 GCTGCTTCTTCTGGGAGCGTGAACGCCGAGCGGAGTCTGCTGCTGAGACCTGT 98817

```


QY	359	GC	GGGGGGGCTCCGGGAGCCTTTGGCTACGAGCGCGGAGTATGTTCATTAGCCCGGTGC	418
Db	98818	GC	GGGGGGGCTCCGGGAGCCTTTGGCTACGAGCGCGGAGTATGTTCATTAGCCCGGTGC	98877
QY	419	CC	AATGCTACGCGCGCGCGCGCAGCGCAACGACCGAGGCGCACACCTTCTCCAGCGCC	478
Db	98878	CC	AATGCTACGCGCGCGCGCGCAGCGCAACGACCGAGGCGCACACCTTCTCCAGCGCC	98937
QY	479	GG	GGTGTTCCGGGCGGGCCCTTCGGAGACCCCACTCTCGCTGGGGGTGGCCTCGGGCT	538
Db	98938	GG	GTGTTCCGGGCGGGCCCTTCGGAGACCCCACTCTCGCTGGGGGTGGCCTCGGGCT	98997
QY	539	GG	ACCCCGGCATCTACGGCCCTTGACCCCTTACAAGCCGCGCGCGGCTTCGGG	598
Db	98998	GG	ACCCCGGCATCTACGGCCCTTGACCCCTTACAAGCCGCGCGGCTTCGGG	99057
QY	599	AG	ATCGTAGCCGCGCAGTCTGGGCGCGCCAAAGCGTTTCGTGCTATCCCGGGTACG	658
Db	99058	AG	ATCGTAGCCGCGCAGTCTGGGCGCGCCAAAGCGTTTCGTGCTATCCCGGGTACG	99117
QY	659	TG	GAGACGCTGGTGGTCGCGGACGAGTCAATGTCAAGTTCACGCGCGCGACCTGGAAC	718
Db	99118	TG	GAGACGCTGGTGGTCGCGGACGAGTCAATGTCAAGTTCACGCGCGCGACCTGGAAC	99177
QY	719	AT	TATCTGTGACGCTGCTGGCAACGCGCGGACACTACCGCCATCCAGCATCTCA	778
Db	99178	AT	TATCTGTGACGCTGCTGGCAACGCGCGGACACTACCGCCATCCAGCATCTCA	99237
QY	779	AC	CCATCAACATGTTGGTCAAGTGTGCTTCTTAGAGATCGTAATCCGGGCCA	838
Db	99238	AC	CCATCAACATGTTGGTCAAGTGTGCTTCTTAGAGATCGTAATCCGGGCCA	99297
QY	839	AG	GTACCGGCAATGCGGGCCCTGACGTGGCGCAACTTCTGTGCTGGCAGAGAAGCTGA	898
Db	99298	AG	GTACCGGCAATGCGGGCCCTGACGTGGCGCAACTTCTGTGCTGGCAGAGAAGCTGA	99357
QY	899	AC	AAAGTGAGTGACAAGCACCCGAGTCTGGGACACTGCATCCTCTTACCAGGCGG	958
Db	99358	AC	AAAGTGAGTGACAAGCACCCGAGTCTGGGACACTGCATCCTCTTACCAGGCGG	99417
QY	959	AC	CTGTGGAGCCACCACCTGTGACAC	996
Db	99418	TG	AGTTGATCTGCGCGTCACCTTTGCACCC	99445

RESULT 7	
AC101990/c	
LOCUS	AC101990 182656 bp DNA linear HTG 21-AUG-2002
DEFINITION	Mus musculus clone RP24-371J2, WORKING DRAFT SEQUENCE, 16 unordered pieces.
ACCESSION	AC101990
VERSION	AC101990.2 GI:22381363
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	house mouse.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
REFERENCE	1 (bases 1 to 182656)
AUTHORS	Birren,B., Nusbaum,C. and Lander,E.
TITLE	Mus musculus, clone RP24-371J2

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 361: contig of 361 bp in length
 * 362 461: gap of 100 bp
 * 462 1538: contig of 1077 bp in length
 * 1539 1638: gap of 100 bp

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
-------	---------	---------

Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17060766.
All repeats were identified using RepeatMasker:
http://frrt.genome.washington.edu/RM/RepeatMasker.htm
Smith, A.F.A. & Green, P. (1996-1997)

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17833
Center clone name: 371_J_2

Center: *Clon* name: *Summary Statistics*
 Sequencing vector: Plasmid, n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 178030 bases at least Q40
 Consensus quality: 179364 bases at least Q30
 Consensus quality: 180647 bases at least Q20
 Insert size: 172000; agarose- ϕ
 Insert size: 181156; sum-of-contigs
 Quality coverage: 8.1 in Q20 bases; agarose- ϕ
 Quality coverage: 7.7 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 361: contig of 361 bp in length
 * 362 461: gap of 100 bp
 * 462 1538: contig of 1077 bp in length
 * 1539 1638: gap of 100 bp

1639	3397:	contig of 1759 bp in length
3398	3497:	gap of 100 bp
3498	4947:	contig of 1450 bp in length
3498	4947:	contig of 1450 bp in length
4948	5047:	gap of 100 bp
5048	6893:	contig of 1846 bp in length
6894	6993:	gap of 100 bp
6994	10395:	contig of 3402 bp in length
10396	10495:	gap of 100 bp
10496	12584:	contig of 2089 bp in length
12585	12684:	gap of 100 bp
12685	14977:	contig of 2293 bp in length
14978	15077:	gap of 100 bp
15078	18181:	contig of 3104 bp in length
18182	18281:	gap of 100 bp
18282	22654:	contig of 4373 bp in length
22655	22754:	gap of 100 bp
22755	31261:	contig of 8507 bp in length
31262	31361:	gap of 100 bp
31362	43587:	contig of 12226 bp in length
43588	43687:	gap of 100 bp
43688	57612:	contig of 13925 bp in length
57613	57712:	gap of 100 bp
57713	77467:	contig of 19755 bp in length
77468	77567:	gap of 100 bp
77568	108724:	contig of 31157 bp in length
108725	108824:	gap of 100 bp
108825	182656:	contig of 73832 bp in length.

[illegible]

Query Match 27.4%; Score 783; DB 2; Length 182656;
 Best Local Similarity 87.3%; Pred. No. 4.5e-131;
 Matches 858; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 1 ATGCTTCTCTGGGCACTCTAACCTGCTTTTCGCGGGGGGAACCGCTGAGGCTGTGAG 60

AC126507
 VERSION AC126507.1 GI:21700463
 KEYWORDS HTG: HTGS_PLAS1
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus

REFERENCE
 AUTHORS

1 (bases 1 to 178764)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaraturge,H.C., Are,J.R., Ayelle,M., Banks,T.,
 Barabara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhaik,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Lucier,R., Luna,P., Ma,J.,
 Lozado,R.J., Lu,X., Lucier,A., Martin,R., Martindale,A., Martinez,E.,
 Maheshwari,M., Mapua,P., Martin,R., Meador,M., Mel,G., Metzker,M.,
 Massey,E., Mawhiney,E., Mcleod,M.P., Morgan,M., Morris,S.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojuben,I., Rolfe,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalobos,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

Direct Submission
 Unpublished
 2 (bases 1 to 178764)
 Worley,K.C.

Direct Submission
 Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 178764)
 Worley,K.C.

Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GZFU
 Center clone name: CH230-254N12
 ----- Summary Statistics
 Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 124573 bases at least Q40
 Consensus quality: 131850 bases at least Q30
 Consensus quality: 137700 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_graft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 49 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1
 1037: contig of 1037 bp in length
 1137: gap of unknown length
 1138: contig of 1113 bp in length
 2251: gap of unknown length
 2351: contig of 1229 bp in length
 3580: gap of unknown length
 3680: contig of 1195 bp in length
 4875: gap of unknown length
 4975: contig of 1551 bp in length
 6525: gap of unknown length
 6626: contig of 1280 bp in length
 7906: gap of unknown length
 8005: contig of 1719 bp in length
 9725: gap of unknown length
 9825: contig of 1618 bp in length
 11443: gap of unknown length
 11543: contig of 1358 bp in length
 12901: gap of unknown length
 13001: contig of 1150 bp in length
 14151: gap of unknown length
 14250: contig of 1315 bp in length
 15565: gap of unknown length
 15666: contig of 1810 bp in length
 17476: gap of unknown length
 17576: contig of 1115 bp in length
 18691: gap of unknown length
 18790: contig of 1479 bp in length
 20369: gap of unknown length
 20370: contig of 1558 bp in length
 21927: gap of unknown length
 21928: contig of 1971 bp in length
 23998: gap of unknown length
 23999: contig of 1581 bp in length
 24099: gap of unknown length
 25680: contig of 1649 bp in length
 25780: gap of unknown length
 27429: contig of 2149 bp in length
 27529: gap of unknown length
 29677: contig of 1100 bp in length
 29778: gap of unknown length
 30878: contig of 1854 bp in length
 30978: gap of unknown length
 32832: gap of unknown length
 32932: contig of 2488 bp in length
 35419: gap of unknown length
 35519: contig of 1874 bp in length
 37393: gap of unknown length
 37494: contig of 2342 bp in length
 38336: gap of unknown length
 39936: contig of 2854 bp in length
 42789: gap of unknown length
 42890: contig of 3757 bp in length
 46646: gap of unknown length
 46747: contig of 2442 bp in length
 49189: gap of unknown length
 49288: contig of 4517 bp in length
 53805: gap of unknown length
 53905: contig of 2297 bp in length
 53906: contig of 2297 bp in length

Tel:81-45-503-9111, Fax:81-45-503-9170)
On Mar 7, 2002 this sequence version replaced gi:13488920.
COMMENT

COMMENT	On Mar 7, 2002 this sequence version replaced gi:13489920.
FEATURES	
source	Location/Qualifiers 1..28000 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="11" /map="11q" /clone="Rp11-211H5"
BASE COUNT	5938 a 6083 c 6679 g 8300 t
ORIGIN	
Query Match	27.2%; Score 777.2; DB 9; Length 28000;
Best Local Similarity	98.4%; Pred. No. 7.2e-130;
Matches	785; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Oy	2056 ACTGGACTTTACCAAGCCCATGATGGCTACAAATTTCGTGTGGCCATCCCGCAGGC 2115
Dd	3450 ACGCCCCCCCTTCCCOCGCCAGGCATGGCTACAAATTTCGTGTGGGCCAATCCCGCAGGC 3509
Oy	2116 GCCTCAAGATCGACATCCGCCACGCCGGTTACAAGAAGGCTGATCGGGGATGACAACACTAC 2175
Dd	3510 GCCTCAAGATCGACATCCGCCACGCCGGTTACAAGAAGGCTGATCGGGGATGACAACACTAC 3569
Oy	2176 CTGGCTCTGAAGAACCAAGCAAGTAAGTACCTGCTCAAGGGGCAATTCGTGGTGCGGC 2235
Dd	3570 CTGGCTCTGAAGAACCAAGCAAGTAAGTACCTGCTCAAGGGGCAATTCGTGGTGCGGC 3629
Oy	2236 GTGGAGCGGCACCTGCTGTGAAGGGCAGTCTGCTGCGGTACAGCGGCACGGGCACAGGC 2295
Dd	3630 GTGGAGCGGCACCTGCTGTGAAGGGCAGTCTGCTGCGGTACAGCGGCACGGGCACAGGC 3689
Oy	2296 GTGGAGAGCTGCAGGCTTCCCGGGCCATCTCTGGAGCCGCTGACCGTGGAGGTCTCTCTCC 2355
Dd	3690 GTGGAGAGCTGCAGGCTTCCCGGGCCATCTCTGGAGCCGCTGACCGTGGAGGTCTCTCTCC 3749
Oy	2356 GTGGGGAAGATGACACCGCCCGGGTCGCGTACTCCTTCTATCTGCCCAAGAGCCTCGG 2415
Dd	3750 GTGGGGAAGATGACACCGCCCGGGTCGCGTACTCCTTCTATCTGCCCAAGAGCCTCGG 3809
Oy	2416 GAGCACAACTCTCTCATCCCAAGGACCCTGCTGTCTGTCTGCACACAGCGTC 2475
Dd	3810 GAGCACAACTCTCTCATCCCAAGGACCCTGCTGTCTGTCTGCACACAGCGTC 3869
Oy	2476 CTCAGCCTCTCCAACAGGTGGAGAGCGGAGACAGGCCCTGTCAGCCTGGGTGGCT 2535
Dd	3870 CTCAGCCTCTCCAACAGGTGGAGAGCGGAGACAGGCCCTGTCAGCCTGGGTGGCT 3929
Oy	2536 GGCAGCTGGGGGCCGTGCTCCCGAGCTGCGCAGTGGCCCTGCAGAAAGCGGGGGTGGAC 2595
Dd	3930 GGCAGCTGGGGGCCGTGCTCCCGAGCTGCGCAGTGGCCCTGCAGAAAGCGGGGGTGGAC 3989
Oy	2596 TGTGGGGCTCCGCCGGGACGCCAGGTCCCTGCTGTGATGAGCCCATCGGCCCGGTG 2655
Dd	3990 TGCCGGGGCTCCGCCGGGACGCCAGGTCCCTGCTGTGATGAGCCCATCGGCCCGGTG 4049
Oy	2656 GAGACAAAGCTCGGGGAGCCCTGCCCACTCGGGAGCTCAGCGCTGTGTACCCCTGC 2715
Dd	4050 GAGACAAAGCTCGGGGAGCCCTGCCCACTCGGGAGCTCAGCGCTGTGTACCCCTGC 4109
Oy	2716 TCCAAGAGCTGGCGCCGGGGATTTCAGAGGCGCTCACTCAAGTGTGGGCCACGGAGGC 2775
Dd	4110 TCCAAGAGCTGGCGCCGGGGATTTCAGAGGCGCTCACTCAAGTGTGTGGGCCACGGAGGC 4169
Oy	2776 CGGTGCTGGCCCGGGACACAGTGCACACTTGCACCAAGCCCCCAGGAGCTGGACTTCTGC 2835
Dd	4170 CGGTGCTGGCCCGGGACACAGTGCACACTTGCACCAAGCCCCCAGGAGCTGGACTTCTGC 4229
Oy	2836 GTCTTAGGGCCGTGCTGA 2853
Dd	4230 GTCTTAGGGCCGTGCTGA 4247

RESULT 10	
BC009667	
LOCUS	
DEFINITION	

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE	AUTHORS	TITLE	JOURNAL
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

REMARK
COMMENT

source

BASE COUNT
ORIGIN

Best Look Matches

Db 13 CAAGTGTAACTAATCTGCCGAGCAATGGGACTGGCTACTCTATGTACTAGCACCTAA 72
QY 1902 GGTGTGTGAGCGGACGCTGTGCTCTCCAGTCTCCACCTCCCTCTGTGTCTAAGCAAGTG 1961
Db 73 GGTGTGTGATGTAGCTGTGTACTCTCTGACTCCACCTCGCTGTGTCTAAGCAAGTG 132
QY 1962 CATCAAGGCTGGCTGTGTGATGGAACTGGGCTCCAGAGAGATTCGACAAGTGTGGGT 2021
Db 133 CATCAAGGCTGGCTGGGAGGAACTGGGCTCCAGAGAGAAATTCGACAATGTGTGT 192
QY 2022 GTGTGGGGAGACAATAAGAGCTGCAAGAGGTGACTGACTCTTCACCAAGCCCATGCA 2081
Db 193 GTGTGGTGGAGACATAGAGCTGTAGAGAGGTGACAGGACTCTTCACCAAGCCCATGCA 252
QY 2082 TGGTACAAATTTGTGTGGGCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGCG 2141
Db 253 CGGCTACAAATTTGTGTGGGCATCCCGTGTAGAGCGGACCTGTGTGTGTGTGTGTGT 312
QY 2142 CGGTTACAAAGGCTGTATCGGGATGACAACTACTCTGGCTCTGAAGACACGCCAAGCAA 2201
Db 313 TGGTTACAAAGGCTGTATCGGGATGACAACTACTCTGGCTCTGAAGACACGCCAAGCAA 372
QY 2202 GTACTGTCTCAACGGGCAATTTGTGTGGTGTGGCGGTGGAGCGACCTGTGTGTGAAGG 2261
Db 373 ATACTGTCTCAATGGGCAATTTGTGTGTATCGCTGTAGAGCGGACCTGTGTGTGAAGG 432
QY 2262 CAGTCTGTGGGTGTATGAGCGGACAGCGGACAGCGGTGGAGAGCTGTGAGGCTTCCCGGCC 2321
Db 433 CAGTGTGTCTATAGTGGCACTGTGTACTGTGAGTGGAGCGCTGTGAGCGCTTCTCGAOC 492
QY 2322 CATCTGAGCGGCTGACCGTGGAGTCTCTCGTGGGGAAGATGACACCGCCCGGCT 2381
Db 493 CATCTGAGCGGCTGACCGTGGAGTCTCTCGTGGGGAAGATGACACCGCCCGGCTGT 552
QY 2382 CGCTACTCTCTATCTGCCAAAGAGCTCGGGAGCAAGTCTCTCTATCCCRAGGA 2441
Db 553 GCGTTATCTCTTACCTGCCAAGAACCTCGGGAGCAAGTCTCTCTATCCCRAGGA 612
QY 2442 CCGCGGGGACCTCTGTCTTGCACAAAGCGTCTCAGGCTCTCCAAACAGTGGAGCA 2501
Db 613 CCGTCAAGGTTCCCTCGTGTGCGCAACAGTGTCTCAGGCTTCTCAACCAAGTAGACA 672
QY 2502 GCGGAGACAGAGCGCCCTGACCGCTGGTGGCTGGGAGCTGGGGCGCTGTCTCCGGAG 2561
Db 673 GCGAGATAACAGACTCTCTGCGGCTGGGTGGAGGAGAGTGGGAGCTTGTCTGGTGTAG 732
QY 2562 CTGCGGAGTGGCTCTGAGAGCGGGCGGTGGAGTGTGGGGCTCGCGCGGAGCGGAC 2621
Db 733 CTGTGGCAGTGGCTCTACAGAGCGAGCAGTGGAGTGGCGGAGCTCCCGAGGAGCAAG 792
QY 2622 GTCCCTGCTCTGATCGAGCGCAATCGGCGGCTGGAGACACAAGCGCTGGCGGAGCGCTG 2681
Db 793 AGCTCTGCTGTGATGTAGACCATCGGCCATTGGAGAGCGAGCGCTGTGGGAACCCCTG 852
QY 2682 CCGCACCTGGAGCTCAGCGCTGTGCTCAGCTCTCCAGAGCTGGGCGTGGGATTTCA 2741
Db 853 CCGAATTTGGAGCTCGGCAACTGTGCGGCTGTCTTAAAGAGTGTGGCGTGGATTTAA 912
QY 2742 GAGCGCTCACTCAAGTGTGTGGGCCACGAGGCGCGGCTGTGCTGCGGAGCAGCTGCA 2801
Db 913 GAGCGCTCCGCTCAAGTGTGTGGGCCACGAGGAGGAGGCTGTGCTGCGGAGCAGCTGTGA 972
QY 2802 CTTCACCGGAGCGGCGGAGGCTGTGCTGCTGAGAGCGGCTGTGA 2853
Db 973 CTTACGCGGAGCGGCGGAGGATTTAGACTTCTCGCTTGTGAGACCGCTGTGA 1024

RESULT 11
E55282
LOCUS E55282 2670 bp DNA linear PAT 31-JAN-2002
DEFINITION Novel metalloprotease and gene of the same.
ACCESSION E55282
VERSION E55282.1 GI:18629795

KEYWORDS JP 2001008687-A/18.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2670)
AUTHORS Yamaji,N., Nishimura,K. and Sasamata,M.
TITLE Novel metalloprotease and gene of the same
JOURNAL Patent: JP 2001008687-A 18 16-JAN-2001;
YAMANOUCHI PHARMACEUT CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2001008687-A/18
PD 16-JAN-2001
PF 25-JUN-1999 JP 1999180973
PR
PI NOBORU YAMAJI, KOICHI NISHIMURA, MIHO SASAMATA
PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC
C12N9/64, C12Q1/37,
PC C12N15/00, C12N5/00
CC
FH Key Location/Qualifiers
FT source 1..2670 /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..2670 /organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 511 a 853 c 867 g 439 t
ORIGIN
Query Match 23.3%; Score 664.4; DB 6; Length 2670;
Best Local Similarity 59.2%; Pred. No. 2e-109;
Matches 1354; Conservative 0; Mismatches 886; Indels 4%; Gaps 11;
QY 157 CTCATTTTTCAGATCACACATTTTCAGAGAGCTTTTACTACCTGACGCGGAGTCT 216
DB 157 CTCGCGCTCCACCTGTCCGCTTCGGCAAGGGCTTCGTCTCGCGCTCGCGCGGACGAC 216
QY 217 CAGTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGGGCTCCCGCTCCAGGGCTCAC 276
DB 217 AGCTTCTCGGCGCGAGTTCAGATTCAGAGCGCTTCGGGGGCTCGCGCGGCGG 273
QY 277 GGGGCTCTTCAGACCTCGGAGCTGCTTCTATTCTGGGAGCTGAACGCCAGCGGAC 336
DB 274 GGGGCGAGCGGGCTCGCGGCTGCTTCTTCGCGCACCGTGAATGGGAGCGCGAG 333
QY 337 TCGTTCGCTGCTGTAGACCTGTGGGGGGCTCCGGGAGCGCTTTGGCTACGAGGCGCC 396
DB 334 TCGTGGCGGCGGTGAGCGCTGTGGCGGGCTGAGCGGCTCTCTCTGTCGACGCGGAG 393
QY 397 GAGTATGTATTAGCCCGCTGCCAATGCTAGCGCGCGGCGGCGCAGCGCAACAGCCAG 456
DB 394 GAGTTCACCATCCAGCGGAGCGGCGGGGCTCTCTGGCTCAGCGCGCGCTCGGAG 453
QY 457 GCGCACACCTTCTCCAGCGCGGGGTGTTCGGGCGGGGCTTCGCGGAGACCGCCATCTCT 516
DB 454 CGCTGGGCTCCCGCGGAGCGCGCCCTCCCGCGAGGACCGAGTGGGAGGTGCGAGCG 513
QY 517 CGCTGGGGGTG-GCTCGGGCTGGAAACCCGCCATCTACGGGCGCTTGACCCCTTACAA 575
DB 514 GGAGAGGGTTCAGAGGAGGAGAGAGAGACCAACAGAGAGAGAGAGAGAGAGAGAGCAA 573
QY 576 GCGCGGGGGGGGCTTCGGGGAGAGTCTGAGCGCGGCGCAGG-----TCTGGGCGCGCC 630
DB 574 GAAGAGGAGGAGAGAGGCTAGCGAGCGCGCCACCGCCCTCGGGGGCCACGAGTAGGACC 633
QY 631 AAGCGTTTCTGTCTATCCCGGCTACCTGGAGAGCGCTGGTGGTGGCGGAGAGTCAATG 690
DB 634 AAGCGTTTGTCTGTAGGCGCGCTTCGTGAAGAGCTGCTGGTGGCGGATGCGCTCAATG 693
QY 691 GTCAGATTCCAGCGCGCGGAGCTTGAACATTTATCTGTGAGCTGTGGCAACGCGGCG 750
DB 694 GCTGCTTCTACGGGGCGGAGCTGCAGAAACCAATCTCTGAGCTTAATGTCTGTGGCAGCC 753

QY 751 CGACTCTACGGCCATCCACGACATCTCTCAACCCATCAATCGTGTGTGTCAGGTGCTG 810
DB 754 CGAATCTACAGACACCCAGGACATCAAGATTCATCACTCATGTGTGTAAGTGTGCTG 813
QY 811 CTCTTTAGAGATGTGACTCCGGGGCCCAAGGTACACGGCAATGCGGCCCTGACGCTGGCC 870
DB 814 ATCGTAGAAGATGAAATATGGGGCCAGAGGTGTCCGCAATATGGGGGCTTACACTGGGT 873
QY 871 AACTCTGTCCCTGGCAGAGAGCTGACAAAGTGAAGTGAAGCAACCCGAGTACTGG 930
DB 874 AACTCTGTCAACTGGCAGGGGCTTTCAACCCAGCCAGCGACCCACCCAGAGCACTAC 933
QY 931 GACACTGCCATCTCTTCAACGAGGAGGACCTGTGTGG---AGCCACCACTCTGTACACC 987
DB 934 GACAGGCCATCTGCTCAACGAGACAACTTCTGTGGCAGGAGGGGCTGTGTACACC 993
QY 988 CTGGGATGCTGATGTGGGTAGCATGTGTGACCCCAAGAGAGCTGTCTCTCATTTGAG 1047
DB 994 CTGGGTGTGACACATCTGGGAGCACTTGTGACCCCAAGAAAGTGTCTCCGTGATCGAG 1053
QY 1048 GAGGATGGGCTTCCATCAGCTTCAACACTGCCACGAGCTGGGCCACGTTTCAACATG 1107
DB 1054 GATGAGGGCTCCAGGGGGCCACCCCTGGCCCATGAACCTAGGCGACGTCCTCAGCATG 1113
QY 1108 CCCCATGACAAATGTGAAGTCTGTGAGGAGGTGTTTGGGAAGCTCCGAGCCCAACACATG 1167
DB 1114 CCCACGACGACTCCAAAGCCCTCACACGCGCTCTCGGGCCCATGGCAAGCACACGCTG 1173
QY 1168 ATGTCCCGACCTCATCCAGATCGAGCGTGGCCACCCCTGTGACGCTGTGAGTGTGCTG 1227
DB 1174 ATGSCACCGCTGTCTGCTCCACCTGAACAGACGCTGCCCTGTGCTCCGCGCCATG 1233
QY 1228 ATCATCACGACTTCTGGACACGGCAGCGTACTGCTCTGTGACCAACCCAGCAAG 1287
DB 1234 TATCTCACAGCTTCTGGACGGGGGACGAGACTGTCTCTGATGCCCTGTGCTGCTG 1293
QY 1288 CCATCTCTCCCTGCCGAGGATGTCGGGGC-----GCCAGCTTACACCTTGAGCCAGCAG 1341
DB 1294 GCCCTGCCCTCCACAGGCGCTCCGGGGCGCATGGCCCTGTACCAAGCTGGACAGCAG 1353
QY 1342 TGGAGTGTGCTTTGGCTGGGCTTCAAGCCCTGTCTTACA-----TGCAGTAC 1392
DB 1354 TGCAGGAGATCTTTGGGGCGGATTTCCGCCACTGCCCAACACCTCTGCTCAGGACGTC 1413
QY 1393 TGCACCAAGCTGTGGTG---CACCGGAAGCCAAAGGACAGATGTGTGCCAGACCCGC 1449
DB 1414 TGGCCCGAGCTTTGGTCCACACTGATGGGGCTGAGCCCTGTGCCACAGGAAGATGGC 1473
QY 1450 CACTTCCCTGGGCGGATGGCAGCAGCTGTGGCGAGGCAAGCTCTCCCTCAAAAGGGCC 1509
DB 1474 AGCCTGCCCTGGGTGACGCGCAGCCCTGGGGCGCTGGGCACTCTGTCTCAGAAGGAGC 1533
QY 1510 TG-----CGTGGAGACACACACTCAACAGCAGAGGTGATGTCTCTGGGCCAA 1563
DB 1534 TGTCTACTGAGGAGGAAGTGGAGAGGCCCAAGCCCTGTGTAGATGGAGCTGGGACCGG 1593
QY 1564 TGGGATCCCTATGGCCCTCTGCTCGGCACATGTGTGGGGCGCTGACAGCTGGCCAGGAGG 1623
DB 1594 TGGGGACCTTGGGAGATGTCTGGACCTGTGGAGGAGGATACAGTTTACACACCT 1653
QY 1624 CAGTGCACCAACCCCTTGCACAGGGGGCAAGTACTGCGAGGAGGTGAGGGTGA 1683
DB 1654 GAGTGCAGAGGACCCGAGCCTCAAGATGGAGGAAGATACTGCTCTGGGTTCGAGAGCCAAG 1713
QY 1684 TACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAG 1743
DB 1714 TACCGATCTGCCACAGGAGGAATGGCCCC-----CCTACGGGAAAGCTTACGGGAG 1767
QY 1744 GAGCAGTGTGAGGCTTTCAACGGGTACACAGCCAGCAGCAACCCGCTCACTCTCGCCGTG 1803
DB 1768 CAGCAGTGTGAGGAATATATGCTACAAATTACACTGACATGGACGGGAATCT---CCTG 1824

QY 1804 GCATGGGTGCCAAGTACTCCGGGGTGTCTCCCGGACAAAGTGCAGCTCATCTGACGA 1863
DB 1825 CAGTGGGTCCCAAGTACTGTGGGGTGTCCCGGACCGCTGCAAGTGTCTTCTGCGGA 1884
QY 1864 GCCAATGSCACTGCTACTTCTATGTCTGGCACCCCAAGTGTGGACGCGCTGTGC 1923
DB 1885 GCCCGGGGAGGAGCGAGTTCAAAGTGTTCGAGGCCAAGTGTATGATGGCACCCCTGTGT 1944
QY 1924 TCTCCTGACTCCACCTCCGCTGTGTCTCAAGCAAGTGCATCAAGGCTGTGTGATGGG 1983
DB 1945 GGGCCAGAACACATGGCCATCTGTGTCCGTGGCCAGTGTCTCAAGGCCGCTGTGACCAT 2004
QY 1984 AACCTGGGCTCCAGAGAGATTCGACAAAGTGTGGGGTGTGTGGGGAGACATAAAGC 2043
DB 2005 GTGTGTGACTCGCCTCGGAAGCTGGCAAAATGCGGGGTGTGTGGGGCAAGGCAATCC 2064
QY 2044 TGCAGAGGTGACTGTGACTCTTCCACCAAGCCCATGCTACAAATTTCTGTGGTGGCC 2103
DB 2065 TGCAGGAGGTCTCCGGTCCCTCACCCCAACCAANTATGTGCTACATGATGATGTCACC 2124
QY 2104 ATCCCGCAGGCGCTCAAGCATCGACATCGCGAGCGGTTACAAAGGCGTGTATGGG 2163
DB 2125 ATCCAGCTGTGTGCCACTAATATTGAGTGAAGCAGCGGAGCCACCGGGTGTGCAAGC 2184
QY 2164 GATGACAACTACTGTGCTCTGAAGAACAGCAAGGCAAGTACCTGCTCAAGCGGCTTTC 2223
DB 2185 GATGGAACTACTGTGGCGCTGAAGACGGCTGATGGGAGTACCTGCTCAAGCGCAACCTG 2244
QY 2224 GTGTGTCCGGGTGGAGCGGAGCTGTGTGTAAGGCGAGTCTGTGCGGTACAGCGGC 2283
DB 2245 GCATCTCTGCCATGAGCAGACATCTGTGTGAAGGGGACCATCTCTGAAGTACAGCGGC 2304
QY 2284 ACGGGACAGCGGTGGAGAGCTCGAGGCTTCCCGGCCATCTCTGGAGCGCTGACCGGT 2343
DB 2305 TCCATGCGCACCTGGAGCGCTCGAGAGCTTCCGGCCCTTGCAGAGAGCCTCTGACAGT 2364
QY 2344 GAGTGTCTCTCGT---GGGGAAGATGACACCGGCCCGGGTCCCTACTCTCTTCTATCIG 2400
DB 2365 CAGCTCTGACAGTCCCTGCGGAGGTCTTCCCCCAAGGTCAAATACACCTTCTTGT 2424
QY 2401 CCCAAGA 2408
DB 2425 CCTAATGA 2432
RESULT 12
E58655 Novel metallic protease.
LOCUS E58655
DEFINITION E58655
ACCESSION E58655.1 GI:18629877
TITLE Novel metallic protease
KEYWORDS JP 2001017183-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2670)
AUTHORS Yamaji,N., Nishimura,K. and Sasamata,M.
TITLE Patent: JP 2001017183-A 3 23-JAN-2001;
JOURNAL YAMANOUCHI PHARMACEUT CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2001017183-A/3
PD 23-JAN-2001
PF 09-JUL-1999 JP 1999196584
PR NOBORU YAMAJI,KOICHI NISHIMURA,MIHO SASAMATA
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
C12N9/50,C12Q1/37,
CC C12N15/00,C12N5/00
FH Key Location/Qualifiers
FT source 1..2670

FEATURES		FT	Location/Qualifiers		/organism="Homo sapiens (human)"	
Source			1..2670			
BASE COUNT			511 a 853 c 867 g 439 t			
ORIGIN						
Query Match			23.3%; Score 664.4; DB 6; Length 2670;			
Best Local Similarity			59.2%; Pred. No. 2e-109;			
Matches 1354; Conservative			0; Mismatches 886; Indels 48; Gaps 11;			
QY	157	CTCATTTTTCAGATCACAGATTCACGAGGACATTTTACCTACACCTGATGCCGATGCT	216			
Db	157	CTCGGGCTCCACCTTCGGCCCTTCGGCAAGGGCTTCGTGCTCGCGCTGGTCCCGAGAC	216			
QY	217	CAGTTCTTGGCTCCCGCTTCCTCACTGACATCTGGCGTCCCTCTCCAGGCTGCTGCTG	276			
Db	217	AGCTTCTTGGCGCGGCTGCGGGCTTCCTTCCTCCGACCGTCAATGGCGGACCCGAG	273			
QY	277	GGGGCTCTTTCAGACCTCGAGCTGCTTCATTTCTGGGGAGCTGACCGTCCGAGCCGAC	336			
Db	274	GGGGCGAGCGGGGCTGCGGGCTGCTTCCTCCGACCGTCAATGGCGGACCCGAG	333			
QY	337	TGCTTCGCTGCTGAGCTGTGGGGGGGCTCCGCGAGGCTTTCGCTACCGAGGCGCC	396			
Db	334	TGCTTGGCGGGGCTGAGCTGTGGCGGGGCTGAGCGCTCTCTTCCTGCTGGAGCGGAG	393			
QY	397	GAGTATGTCATTAGCCCTCCCAATGCTAGCGCGCGGGGCTTCGGAGATCCCACTCT	456			
Db	394	GAGTACCATCCAGCCGCGGGGCTCCCTCCGCGAGGACCCGAGTGGGAGTGGAGCG	453			
QY	457	GGCGCACACCTTCCAGCGCGGGGTTCGGGGCGGGCTTCGGAGATCCCACTCT	516			
Db	454	CGCTGGGTCCCGCGGAGCGCGCCCTCCGCGAGGACCCGAGTGGGAGTGGAGCG	513			
QY	517	CGCTCGGGGTG-GCTCGGGCTGGAACCCCGCCATCTTCAGGGCCCTGACCCCTTACAA	575			
Db	514	GGAGGGGTGAGAGGAGGAGGAGACCCAGGAGGACACCGAGGAGGAGGAGCA	573			
QY	576	GGCGGGCGGGGCTTCGGGGAGAGTCTGAGCGCGGAGG- - - - -TCTGGCGCGCC	630			
Db	574	GAAGAGGAGGAGAGGGCTAGCGCGCGCCACCGCCCTTCGGGGCCACAGTAGGACC	633			
QY	631	AAGCGTTTCGTCTATCCCGCGTACGTGGAGAGCTGTGTGTCGGCGAGCTCAATG	690			
Db	634	AAGCGGTTTGTCTGAGCGCGCTTCGTGAAGACGCTGCTGTGGTGGCGATGCTCCATG	693			
QY	691	GTCAGTTCACGCGCGGACCTGMACTTATCTGCTGAGCTGTGGTGAACGGCGCG	750			
Db	694	GCTGCTTCTACGGGGCGGCTGCAGAACCCATCTGAGCTTAATGTCTGTGGAGCC	753			
QY	751	CGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCTGTGTGTCAGAGTCTG	810			
Db	754	CGAATCTACAGCACCCAGCATCAAGATTCATCAACTGATGTGTGTTAAAGTCTG	813			
QY	811	CTTCTTAGAGATCTGATCCGGGCCAAGTCCACGGCAATGGGCCCTGACGCTGCGC	870			
Db	814	ATCGTAGAAGATGAAATGGGGCCAGAGTGTCCGACAAATGGGGGCTTACACTGCGT	873			
QY	871	AACCTTCTGTGCTGGCAGAGAGCTCAACAAGTGTGACAGACACCGTGGAGTCTGG	930			
Db	874	AACCTTCTGCACTGGCAGCGGGTTTCAACCGCCAGCGACCCGACCTAGAGCACTAC	933			
QY	931	GACACTGCCATCTCTTACAGGAGGAGCTGTGTGG- - - - -AGCCACCACTCTGACACC	987			
Db	934	GACAGGCCATCTCTGCTCACACAGAACTTCTGTGGGAGGAGGGGCTGTGTGACACC	993			
QY	988	CTGGGATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGCTGTGTCATTGAG	1047			
Db	994	CTGGGTGGCAGACATCGGGGACCATTTGTGACCCCAACAAAGCTGCTCTCGTATCGAG	1053			
QY	1048	GACGATGGGCTTCATCAGCTTCACCACTGCCACAGCTGGGCGCACCGTGTTCAACATG	1107			
Db	1054	GATAGGGGCTCCAGGCGGCCACACCTGGCCCATAACTAGGGCACGCTCTCAAGCATG	1113			
QY	1108	CCCATGACAAATGTGAAAGTCTGTGAGAGGTGTTTGGGAAGCTCCGAGCAACACCATG	1167			
Db	1114	CCCAACGACGACTCAAGCCCTGCACACGAGCTCTTTCGGGCCCATGGCAAGCAGGCTG	1173			
QY	1168	ATGTCCCGGACCTTCATCCAGATCGACCGTGCACACCCCTGGTTCAGCTGCTGCTGCC	1227			
Db	1174	ATGGCACCGGTGTTCGTCCACCTGACACGAGCGTGCCTTGGTCCCTCGAGCGCATG	1233			
QY	1228	ATCATCACCGACTTCTCTGGACAGCGGCTGACTGCTCTGCTGCTGCTGCTGCTGCTG	1287			
Db	1234	TATCTCACAGAGTCTTGGACGGGCGGACGAGACTGTCTCTGATGCTGCTGCTGCTG	1293			
QY	1288	CCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1341			
Db	1294	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1353			
QY	1342	TGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1392			
Db	1354	TGAGGAGAGTCTTTCGGGCGGATTTCCGCACTGCTGCTGCTGCTGCTGCTGCTG	1413			
QY	1393	TGACCAAGCTGTGGTG- - - - -CACCGGAGGCGCAAGGACAGATGCTGCTGCTGCTG	1449			
Db	1414	TGCGCCAGCTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1473			
QY	1450	CACCTTCCCTGGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1509			
Db	1474	AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1533			
QY	1510	TG- - - - -CGTGGAGAGACACACCTCAACAGCACAGGCTGATGCTTCTGCTGCTG	1563			
Db	1534	TGCTACCTGAGGAGAGTGGAGGCGCCAGCCGCTGCTGCTGCTGCTGCTGCTGCTG	1593			
QY	1564	TGGATTCCTTATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1623			
Db	1594	TGGGACCTTGGGAGAAATGTTCTCGGACCTGTGGAGGAGGATACAGTTTTCACACCT	1653			
QY	1624	CAGTGACCAACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1683			
Db	1654	GAGTCAAGAGACCCCGAGCTTCAGAAATGAGGAAGATGCTGCTGCTGCTGCTGCTG	1713			
QY	1684	TACCGATCTGCAATCTGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1743			
Db	1714	TACCACTCATGCCACAGGAGGATGCCCC- - - - -CCTGACGGGAAAGCTTCAGGGAG	1767			
QY	1744	GAGCAGTGTGAGGCTTTCAACGGCTACACACAGCAGCAGCAGCAGCAGCAGCAGC	1803			
Db	1768	CAGCAGTGTGAGAAATTAATGCTACAATTACACTGACATGAGCGGAACTCT- - -CCTG	1824			
QY	1804	GCATGGTGCACCAAGTACTCCGCGCTGCTCCCGGAGCAAGTGCAGCTCATCTCCGA	1863			
Db	1825	CAGTGGGTCCCCAAGTATGCTGGGTGCTCCCGGAGCTGCTGCTGCTGCTGCTGCTG	1884			
QY	1864	GCAATGGCACTGGCTTACTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1923			
Db	1885	GCCCGGGGAGGAGGAGTCAAGTGTTCAGGCGCAAGGATGATGATGGCACTGCTGCTG	1944			
QY	1924	TCCTGACTCCACCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1983			
Db	1945	GGCCAGAAACACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2004			
QY	1984	AACCTGGGCTCCAGAAAGATTCACAGTGTGGGTGCTGCTGCTGCTGCTGCTGCTG	2043			
Db	2005	GTGGTGGCTGCGCTGCGGAGCTGGGAAATGCGGGGTGCTGCTGCTGCTGCTGCTG	2064			
QY	2044	TGCAAGAAGTGAAGTCTTTCACAGGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTG	2103			
Db	2065	TGCAAGAAGTCTGCGGGTCCCTCACCCCGGCAATTTGCTGCTGCTGCTGCTGCTG	2124			
QY	2104	ATCCCGGAGGCGCTCAAGCATTCAGATTCGCTGCTGCTGCTGCTGCTGCTGCTG	2163			

Db 2125 ATCCAGCTGGTCCCACTAATATGTAGCTGAAGCAGCGGACCCCGGCTGTGCAGAAC 2184

Qy 2164 GATGACAACTACCTGCTCTGAAGAACAGCCAGGCAAGTACTGCTCAACGGGCAATTTC 2223

Db 2185 GATGGGAACCTACCTGGCGCTGAAGAGCGGTGATGGCGAGTACTGCTCAACGGCACTTG 2244

Qy 2224 GTGCTGTGCGGCGGTGGAGGGGACCTGGTGGTGAAGGGCAGTCTGCTCGCGGTACAGCGGC 2283

Db 2245 GCCATCTCTGCCATAGCAGGACATCTTGGTGAAGGGACCATCTCTGAAGTACAGCGGC 2304

Qy 2284 ACGGGACACGGGTGAGAGCCTGAGGCTTCCCGGCCCATCTCTGAGAGCGGTGACCGTG 2343

Db 2305 TCCATGCCACCTGAGCGCTGAGAGCTTCCGGCCCTTGCAGAGCCTCTGACAGTG 2364

Qy 2344 GAGTCTCTCCCT---GGGGAAGATGACACGCCGCCGGTCCGCTACTCTCTATCTG 2400

Db 2365 CAGCTCTGACAGTCCCTGGGAGGTCTTCCCGCCCAAAAGTCAATATACACTCTCTTTGT 2424

Qy 2401 CCCAAAGA 2408

Db 2425 CCTAATGA 2432

RESULT 13

AF060153

LOCUS

DEFINITION Homo sapiens METH2 protein (METH2) mRNA, complete cds.

ACCESSION AF060153

VERSION AF060153.1 GI:5725507

KEYWORDS

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3711)

Vazquez,F., Hastings,G., Ortega,M.A., Lane,T.F., Oikemus,S., Lombardo,M., and Iruela-Arispe,M.L.

METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity

J. Biol. Chem. 274 (33), 23349-23357 (1999)

9367466

10438512

2 (bases 1 to 3711)

Vazquez,F., Hastings,G., Ortega,M.-A., Lane,T.F., Lombardo,M., Oikemus,S. and Iruela-Arispe,M.L.

Direct Submission

Submitted (16-APR-1998) Pathology, Beth Israel Deaconess Medical Center, 99, Brookline Avenue, Boston, MA 02214, USA

Location/Qualifiers

1. .3711

/organism="Homo sapiens"

/db_xref="taxon:9606"

1. .3711

/gene="METH2"

706..3378

/gene="METH2"

/note="extracellular protein with metalloprotease and TSP domains"

/codon_start=1

/product="METH2 protein"

/protein_id="A048081.1"

/db_xref="GI:5725508"

/translation="MPPAAPRWLFLLLLLLLPLARGAPARPAAGGQASLVVP

TRLPAGAGELALHLSFGKGVFLRLAPDDSFAPFKIERLGGSGRATGERLGRGCF

FSGTVNGPELSALSLRGLSGFLDGEFTIQPGAGGSLAQPHRLQWRGAPGAR

PLRPGWEVETGQQRQERDHOEDSEESQEEAEGASEPPPLPGATSRKRVSE

ARVETLLVADASMAAFYADLONHILLTMSVAARTYKHPISIKNSINLMVAVLVIED

EKNQPEVSDNGILLNFCNWRFRNPQSDRPEHYDTAILLTRONFCQEGELDTLG

VADIGTICPNKSCVSEDEGLQAHLAHLGHLVLSMHDSDSKPTRLFGPMGRHHV

MAFLFHLNLTLPWSCVSAWYLFELDGGHGDCLDAPGAALPLPTGLPGRMALYQLD

QOCRLQIFDFRCPISATQVCAQLWCHTGAEPCHTKNGSLPWADGTGCGHLIC

SESGCLPEEERPKVPDVGWAPWPGWGECSRTCGGVQVFSHRECKDPENGGRYC

LGRRAVQSCHTECPDGKSPFEQCEKYNAVNYTMDGNLLQWPKVAGYSPDRDC

KLFCRARGRSEFKVFEAKVIDGTLCGPETLAIICVRQCQKAGCDHVDSPRKLDKGV

CGKGNCSRKVSGLPTTNYGNDIVITIPAGATNIDVOKRSPGVQNDGNILAKTAD

GGKLLGNLAISIEQDILVKGITILKYSGISIATLERQSPFLPTLVOLLTPGEV

FPPKVKYTFVPNDVDFSMOSSKERATTNLIQPLHAQWLVGDWSECSSTCGAGWRR

TVCEKRPSCGASATCNKALPKEDAKPESQLCL"

BASE COUNT 733 a 1126 c 1201 g 646 t 5 others

ORIGIN

Query Match 23.3%; Score 664.4; DB 9; Length 3711;

Best Local Similarity 59.2%; Pred. No. 1.9e-109;

Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

Qy 157 CTCATTTTTCAGATCACAGCATTTTCAGAGGACGTTTTCCTTACACCTGACGCGCGATGCT 216

Db 865 CTCGGCTCCACCTGTCCGCCCTTCGGCAAGGCTTCGTGTTCGCCCTCGGCGCCGACGAC 924

Qy 217 CAGTCTTGGCTCCGCCCTTCCACTGAGCATCTGGSGCTGCCCTCCAGGGCTCAC 276

Db 925 AGCTTCTGGCGCCCGAGTTCAGATCGAGCGCTCGGGGCTCCGGCGGGCG---ACC 981

Qy 277 GGGGCTCTTCAGACCTCGACGCTGCTTCTATTCTGGGGACGTGAACGCGGACGCCGAC 336

Db 982 GGGGCGAGCGGGGCTCGCGCGCTGTTTTTTCGGCACCGTGAATGGGAGGCCCGAG 1041

Qy 337 TCGTTCGCTGTGTGAGCCTGTGCGGGGGCTTCGCGGAGGCTTTGGCTACCGAGCGCC 396

Db 1042 TCGCTGGCGGGGTGAGCCTGTGCGGGGCTGAGCGCTCTCTTCCTGCTGGAGCGGGAG 1101

Qy 397 GAGTATGTCTTAGCGCGCTGCCAATCTAGCGCGCGCGGCGGACGCAACAGCGAC 456

Db 1102 GAGTTCACCATCCAGCCGAGCGCGGGGGCTCCCTGGCTCAGCGCACCGCTGCAG 1161

Qy 457 GCGCACACCTTCTCCAGCGCGGGGGTGTTCGGCGGGGCTTCGCGAGACCCACCTCT 516

Db 1162 CGCTGGGTCCGCGGAGCGCGCCCTCCCGGAGGACCCGAGTGGGAGGTGGAGACG 1221

Qy 517 CGCTCGGGGTG-GCTCGGCTGAACCCCGCATCTACGGGCGCTTGACCCCTTACAA 575

Db 1222 GGAGAGGTCAGAGGACGAGAGGAGGACCAACAGGAGGACGAGGAGGAGGAGAGUCAA 1281

Qy 576 GCGCGCGGGCGGGCTTCGGGGAGAGTCGTAGCGCGCGCGCAGG-----TCTGGCGCGCC 630

Db 1282 GAAGAGGAGGAGAGGCGCTAGCGAGCGCGCCACCGCCCTGGGGCCACGAGTAGGACC 1341

Qy 631 AAGCGTTTCGTCTCTATCCCGGGTACGTGAGAGCGCTGTGTCGCGGAGCAGTCAATG 690

Db 1342 AAGCGTTTGTCTGTGAGCGCGCTTCGTGAGAGCGTGTGTTGGCGGATGCTGTCCAIG 1401

Qy 691 GTCAAGTTCACGCGCGGACCTGGAACATTATCTGTGACGCTGCTGGCAACGCGGCG 750

Db 1402 GTGCTCTTCTACGGGGCGGACCTGCAGAACACATCTCTGAGCTTAATCTCTGTGGACCC 1461

Qy 751 CGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCGTTTGGTCAAGTGTG 810

Db 1462 CGAATCTACAAGCACCCAGCATCAAGAATTCATCAACCTGATGGTGTAAAAGTGTG 1521

Qy 811 CTTCTTAGATCTGATCTCCGGGCGGAGGTCACCGCAATGCGGCTTCACGCTGCGC 870

Db 1522 ATCTGTAGAAGATGAAAAATGGGGCGGAGGTGTCGGAACATGGGGGCTTACACTGCGT 1581

Qy 871 AACTTCTGTGCTGGCAGAAAGCTGAACAAAGTGAAGTGAACAGCACCCGAGTACTGG 930

Db 1582 AACTTCTGAACTGGCAGGGGGTTCACACGCGCCAGCGACCCGAGGACTAC 1641

Qy 931 GACACTGCCATCCTTCTACAGGAGGACCTGTGTGG---AGCCACCACTGTGACACC 987

Db 1642 GACACGGCCATCCTGTCTACCAAGACAACTTCTGTGGGAGGAGGGGCTGTGTGACACC 1701

Qy 988 CTGGGCTGCTGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047

Db 1702 CTGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1761

Qy 1048 GACGATGGCTTCCATCATGCTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107

[illegible]

Db	2833	ATCCAGCTGGTGCCACTAATATTGACGTGAAGCAGCGGAGCCACCCGGGTGTGCAGAAC	2892
Qy	2164	GATGACAACTACCTGGCTCTGAAGAACAGCGCAAGTACCTGCTCAACGGCANTTTC	2223
Db	2893	GATGGAACTACCTGGCGCTGAGACGGCTGATGGCAGTACCTGCTCAACGGCAACTG	2952
Qy	2224	GTGGTGTGGCGGGTGAACGGGACCTGGTGGTGAAGGSCAGTCTGCTGCGGTACAGCGCC	2283
Db	2953	GCCATCTCTGCATAGACGAGCATCTGTGTGAAGGGGACCATCTTGAAGTACAGCGGC	3012
Qy	2284	ACGGGCACAGGGTGGAGAGCCTCGAGCTTCCCGGCCCATCTTGAGAGCCCTGACCGTG	2343
Db	3013	TCCATFCGCCACCTGGAGCGCTCGAGAGCTTCGCGGCTTGCAGAGACCTCTGACAGTG	3072
Qy	2344	GAGTCTCTCTCCGT---GGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTATCTG	2400
Db	3073	CAGTCTCTGACAGTCCCTGGCGAGGTCTTCCGCCAAAGTCAATACACCTCTTTGT	3132
Qy	2401	CCCAAGA 2408	
Db	3133	CCTAATGA 3140	
RESULT	14		
LOCUS	D67076		
DEFINITION	Mouse mRNA for secretory protein containing thrombospondin motifs, complete cds.		
ACCESSION	D67076	4180 bp	mRNA linear
KEYWORDS	D67076.1 GI:1813339		ROD 02-FEB-1999
SOURCE	secretory protein containing thrombospondin motifs; ADAMTS-1. Mus musculus colon adenocarcinoma cell_line:murine colon 26 cdna to mRNA.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (sites)		
TITLE	Kuno,K., Kanada,N., Nakashima,E., Fujiki,F., Ichimura,F. and Matsushima,K.		
JOURNAL	Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene		
MEDLINE	J. Biol. Chem. 272 (1), 556-562 (1997)		
REFERENCE	97150761		
AUTHORS	2 (bases 1 to 4180)		
TITLE	Kuno,K., Kanada,N., Nakashima,E., Fujiki,F., Ichimura,F. and Matsushima,K.		
JOURNAL	Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene		
REFERENCE	Unpublished		
AUTHORS	3 (bases 1 to 4180)		
TITLE	Kuno,K.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (28-SEP-1995) Kouji Kuno, Cancer Research Institute, Kanazawa University, Pharmacology; Takara-machi 13-1, Kanazawa, Ishikawa 920, Japan (Tel:0762-62-8151(ex.5454), Fax:0762-60-7704)		
FEATURES	Location/Qualifiers		
SOURCE	1..4180		
gene	/organism="Mus musculus"		
CDS	/db_xref="taxon:10090"		
	/cell_line="murine colon 26"		
	/cell_type="colon adenocarcinoma"		
	1..4180		
	/gene="ADAMTS-1"		
	441..3296		
	/gene="ADAMTS-1"		
	/note="putative"		
	/codon_start=1		
	/product="secretory protein containing thrombospondin motifs"		
	/protein_id="BAAL1088.1"		
	/db_xref="GI:1813340"		

/translation="MGDVORAAARSGLSAHMLLLLLLSTMLLCARGHRPTEEDE
ELVPLSRAPCHDSTTLRLDLDAFG00LHLKLPDSGE LAPGFTLOTVGRSPGSEAO
HLDPTGDLNARCVSTVNGDPSAALSILCEGRCFAFLQGEPEFIOPAGVATERLA
PAVPEESARCFHLLRRRGSGKACGVMDDELPLI SDSRPESQNRQNPVRDP
TPDQAKSPGSGIRKRVVSEPRYETMLVADQSMADFHSGSLKHYLLTLFVSAARF
YKPSIRNSISLVVKKILVIEQKPEVTNAAITLRNFCPSWKQHNPSRDPDEHY
DTAILFTRODLGSHSTCDTFLGMADYVTCDFPSRSCSVIEDGLQAAFTAHLEGHVEN
MPHDAKHCSLNGSGDHLMSLSLSDHSPSPGSAVMVTSFLDNHGECLMDK
PQNPILKPSDLPTLYLDANROCOFTFGEESKHCDDAASTCTLACTGTSGLLVCOTK
HPMDWGTSGGKKWCYSGKVNKTDMDKHFATPVHGSWGPWPGWDCSRTGGGVQYT
MRECONPVKNGCEGRKRYRSCNLEDPCDNNKTFREEQCAHNEFSAKSGNE
PIVETWPKYSPKDRCKLCEAKGTFYFVLQPKVDGTPCSPDSVSCVQGCQVK
AGCDRLIDSKKFGDCCGGGSGCKKMSGIVTSTRPGYHDIVIPAGANIEVKR
NORGSNNSSFLAIRAADQYILNGNFTLSTLEODLYTKVLRVSGSSAALERTSRF
SPLKEPLTLOVLVGHGALRPKIKFTYEMKKKTESFNALPTFSEWIEBWSCKTCGS
GWQRRVQCRDINGHPASCAKEVPKASTRPCADLPCPHWQVGDWSPCKTCGKGYKK
RLKCVSHDGGVLSNESCDPLKPKHYHIDFCTLTQCS"
BASE COUNT 1078 a 1051 c 1146 g 905 t
ORIGIN

Query Match 22.2%; Score 634.4; DB 10; Length 4180;
Best Local Similarity 60.6%; Pred. No. 4.5e-104;
Matches 1097; Conservative 0; Mismatches 701; Indels 12; Gaps 3;

Qy	631	AAGCGTTTCGTGTATFCCCGCGGTAGTGGAGACGCTGTGTGTCGCGGACGAGTCAATG	690
Db	1143	AAGCGATTTGTGTCACGCCCGCTTATGTGGAACACCTGCTAGTACACCATCATG	1202
Qy	691	GTCAGTTTCCACGGCGGACCTTGAACATTTATCTGCTGACGCTGCTGTCGACGGCGG	750
Db	1203	GCGGACTTCCACGGCAGCGCTCTAAAGCATTTACCTTCTAAACCTGTCTCTGTTGGG	1262
Qy	751	CGACTTACCGCCATCCACGATCCTCAACCCCATCAACATCTGTGTGTTCAAGTGTCTG	810
Db	1263	AGTTTTACAGCATCCACGATAGGAATTAATAGCTGTGTGTGTTGAGATCTTG	1322
Qy	811	CTTCTTAGAGATCGTACTCGGCGCCCAAGTCAACCGCAATCGGCGCTGACGCTGCGC	870
Db	1323	GTCATATACGAGGACGAGAGGACCAAGAGTTACCTTCCAAATGACGCTCTCACCCCTCG	1382
Qy	871	AACTTCTGCTGGCAGAGAGAGCTCAACAAAGTGTGACAGACGACCCGAGTACTGG	930
Db	1383	AAATTTGCGAGCTGGCAGAAACACACAGCCCGAGTACCGCGGATCCAGAGCACTAT	1442
Qy	931	GACACTGCCATCTCTTACCAGGACGAGCTGTGTGGAGCCCAACCTGTGACACCCCTG	990
Db	1443	GACACTGCAATTTCTTTCACAGACAGGATTTATGTGGCTCCCAACAGCTGTGACACTC	1502
Qy	991	GGCATGGCTGATGTGGTACCATGTGTGACCCCAAGAGAGCTGCTGTGATGTAGGAC	1050
Db	1503	GGAATGGCAGATGTGGAACCGTATGTGACCCAGCAGAGAGCTGCTCAGTATAGAAGT	1562
Qy	1051	GATGGCTTCCATCAGCTTCAACACTGCCACGAGCTGGGCCACGTGTTCACACATGCC	1110
Db	1563	GATGGTTTGAAGCGCTTCCACAGAGCCCATGAATTTGGGCCATGTGTTTACATGCGG	1622
Qy	1111	CATGACAAATGTAAAGTGTGTGAGAGGTTTGGGAAGTTCGAGCCCAACACATGATG	1170
Db	1623	CACGATGATGTAAAGCACTGTGCCAGCTTGAATGGTGTGAGTGGCGATTCTCATCTG	1682
Qy	1171	TCCCGGACCTCATCAGATCGACCGTGCACACCCCTGTCAGCTGCGTGGCATC	1230
Db	1683	GCTTCGATGCTCTCCAGTTAGACATAGCCAGCCCTGGTACCTTTCGATGTGCTACATG	1742
Qy	1231	ATCACCGACTTCTTCGAGCGGCGAGGTGACTGCTCCTTGGACCAACCCAGCAAGCCC	1290
Db	1743	GTCAGCTCTTCTTATATGAGACACGGGAATGTTTGTGAGCAAGCCCGAGATCCA	1802
Qy	1291	ATCTCCCTGCCGAGGATCTGCCGGGCGGACGCTACACCTGACCCAGCAGTGGAGCTG	1350
Db	1803	ATCAAGCTCCCTTCTGTATCTTCCCGGTACCTTGTACGATGCCAACCCGACGTGCTAG	1862
Qy	1351	GCTTTTGGCGTGGCTCCAGCCCTGCTCTTATACATGACGATAC ---TGCAACAAAGCTGTG	1407

Db	1863	ACATTCGGAGAGGAATCCAAAGCACTGCGCTGATGAGCCAGCACATGTACTACCTCTGG	1922
Qy	1408	TGCACCGGAAGGCCAAGGGACAGATGTTGCCAGACCCGCCACTTCCCTGTCGGCGAT	1467
Db	1923	TGCATGCGACCTCCGGTGGCTTACTGGTGTGCCAAACAAACACTTCCCTTGGGCGAT	1982
Qy	1468	GGCACCACTGTGGCGAGGGCAAGCTTGCCTCAAGGGCGCTCGTGGAGAGACACAAC	1527
Db	1983	GGCACCACTCTGGAGAGGAAGTGTGTCTCAGTGGCAAGTCCGTGAACAACAGACAG	2042
Qy	1528	CTCAACAAGACAG-----GGTGGATGTTCTTGGGCCAATGGATCCCTATGGCCCC	1581
Db	2043	ATGAAGCAATTTTGTACTCTCTTTCATGGAAGCTGGGGACCATGGGACCGTGGGAGAC	2102
Qy	1582	TGCTCGCCACATGTGGTGGGGGTGCAGCTGGCCAGGAGGAGTGCACCAACCCACCC	1641
Db	2103	TGCTCAAGAACTGTGGTGGTGGATTCATACACATGAGAGATGTGACCAACCCAGTC	2162
Qy	1642	CCTGCCAACGGGGCAAGTACTGCGAGGGAGTGGGTGAAATACCGATCTTCAATCTG	1701
Db	2163	CCAAAGAACGGGGAAGTACTGTGAAGCAACAGAGTCCGCTACAGTCTCTGTAACAIC	2222
Qy	1702	GAGCCTCCGCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGGCTTC	1761
Db	2223	GAGGACTGTCCAGACAATAA---CGGAAAAACGTTACAGAGAGGAGTGTGCGGCGCAC	2279
Qy	1762	AACGCTTACAACACAGCACCAACCGGCTCCTCTCGCGGTGGCATGGGTGCCCAAGTAC	1821
Db	2280	AATGAGTTTTCCAAAGCTTCTTTGGGAATGAGCCCACTGTAGAGTGGACACCCAGTAC	2339
Qy	1822	TCCGCGGTGTCTCCCGGACAAAGTCAAGTCAATCTATCTCCGAGCCATGGCACTGGCTAC	1881
Db	2340	GCGCGGTCTCGCCAAAGGACAGGTGCAAGCTACCTGTGAAGCCAAAGGATTTGGTAC	2399
Qy	1882	TCTATGTGCTGGCACCAACAGGTGTGACGCGAGCTGCTCTCTGACTCCACCTCC	1941
Db	2400	TTTTTGTGCTTACAGCCCAAGGTGTAGTGGACTCCCTGTAGTCCAGACTTACCTCT	2459
Qy	1942	GTCTGTGTCCAGGCAAGTGCATCAAGCTGGCTGTGTATGGGAACTGGGTGCCAAGAAG	2001
Db	2460	GTCTGTGTCAAGGCGAGTGTGTGAAGCTGGCTGTGTATCGATCATAGACTTCCAAAAA	2519
Qy	2002	AGATTTCAGAGTGTGGGTGTGGGGGAGACAATAGAGCTGCAAGAGGTGACTGGA	2061
Db	2520	AAAGTTGATAGTGTGGGCTTTGTGGAGGAAACGGTTCCACATGCAAGAGATGTCAGGA	2579
Qy	2062	CTCTTCAACCAAGCCCATGCTAGCTACAAATTTGCTGGTGGCCATCCCGCAGGCGCTCA	2121
Db	2580	ATAGTCACTAGTACAAGACCTGGGTATCATGACATTTGTCACAATTTCTGCTGGAGCCAC	2639
Qy	2122	AGCATCGACATCCGCCAGCGGTTTACAAAGGCTGTATGAGGCTGCGGATGACAATCTGGCT	2181
Db	2640	AACATTTGAAGTGAACATCGGAATCAAAAGGGGTCCAGAAAACATGGCAGCTTTTGGCT	2699
Qy	2182	CTGAAGACAGCAAGGCAAGTACCTGCTCAACGGGCAATTTCTGGTGTGCGCGGTGGAG	2241
Db	2700	ATTAGAGCGCTGTAGTGTACCTTATCTGATGGAATCACTTCTCTCCACACTAGAG	2759
Qy	2242	CGGGACCTGGTGGTGAAGGGCAGTCTGCTCGGTACAGCGGACCGGACACGCGTGGAG	2301
Db	2760	CAAGACCTCAOCTACAAAAGTACTGCTTAAGGTACAGTGGTTCTCGCTGCGCTGGAA	2819
Qy	2302	AGCTGACGCTTCCGGCCCATCTCTGGAGCGGCTGACCGTGGAGGTCTCTCCCTGGGG	2361
Db	2820	AGAAATCCGAGCTTTAGTCCACTCAAGAACCCCTTAACCATCCAGGTTCTTATGTTAGGC	2879
Qy	2362	AAAGTACACCCCGGCTCGCTACTCTTCTATCTGCCCCAAAGAGCTCGGGAGGAC	2421
Db	2880	CATGCTCTCCGACCCCAAAATTAATTTACCTACTTTATGAAGAACAGACAGTCAATC	2939
Qy	2422	AAAGTCTCTC 2431	

Db	2940	AAGCCATTC	2949
RESULT 15	AF207664	4659 bp	mRNA linear PRI 10-JAN-2000
LOCUS	Homo sapiens matrix metalloprotease (ADAMTS1) mRNA, complete cds.		
ACCESSION	AF207664		
VERSION	AF207664.1	GI:6685071	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Glienke, J., Schmitt, A., Pillarsky, C., Hinzmann, B., Weiss, B., Rosenthal, A., and Thierauch, K. H.		
AUTHORS	Glienke, J., Schmitt, A., Pillarsky, C., Hinzmann, B., Weiss, B., Rosenthal, A., and Thierauch, K. H.		
TITLE	Genes differentially expressed by endothelial cells in distinct angiogenic states		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 4659)		
AUTHORS	Glienke, J., Schmitt, A., Pillarsky, C., Hinzmann, B., Weiss, B., Rosenthal, A., and Thierauch, K. H.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-NOV-1999) Experimental Oncology, Schering AG, Muellerstr 178, Berlin 13342, Germany		
FEATURES	Location/Qualifiers		
source	1..4659		
gene	/organism="Homo sapiens"		
5'UTR	/db_xref="taxon:9606"		
CDS	/cell_type="endothelial cells"		
	1..4659		
	1..442		
	443..3346		
	/gene="ADAMTS1"		
	/note="METH-1-like protein"		
	/product="matrix metalloprotease"		
	/protein_id="AAF23772.1"		
	/db_xref="GI:6685072"		
	/translation="MQRAVPGFGRRKGLSDMGNAERAPGSRSGPVPVPTLLLAALL		
	AVSDALGPSDEDELLVPELERAPGHGTTIRLHFDDQDLDELDPSSFLAPFTPL		
	QNVKRSSETPLEPTDLAHGFCYFTVNGDPSAAALSLGCEVRGAFYLLGEAVFIPL		
	LPAAERLTAAPGKGPAPLQFHLRRNRQDVGCTGCVVDDPRPTGAETDEDE		
	GTEGEDAAGSPDPAQGVQPTGTSIRKRFVSSHRYVETMLVADDSMAEFHGS		
	GLKHYLLFLFVAARLYKPSIRNSLVVVKILVHDEQKPEVTSNAALTILNFEN		
	WKOANPPSDRDAEHYDPAILFTQDLCGSOTDITLGHADVTGCDPSRSCVIEDDG		
	IQAAFTTAHELGVLMPPHDDAKOCSALNGVNDQSHMASLMSLDHSCPDASTST		
	ITSTFDNGHCELMKDPQNTOLPGLDLPSTSYDANROCFTEGDSKHCPSAAYM		
	LMWCTSGVLVQCTQKHPWADGTSCGPKWCINGKVNKTDRKFDTPFGSGMWG		
	PWGDSTRCTGVVQYTHRECDNVPKNGKYCEKRYRSCNLEDCPDNNGKTERE		
	QCEAHNEFKASFGSPAVEMIPKAVGSPKDRCKLQAKGTYGFLVLPKVDGTP		
	CPSDSTSVCOGQVKAQDRILIDSKKFKDKGCVGGNGSTCKRISGSVTSAPGVHD		
	ILTIPTGATNIEVQRNQRSGNSFLAKAAGDTYILNGDITSLFLEDQIMYGVV		
	LYKSSAALERIRSFSPLEPLTIQVLTGVNLRPKIYTYFVKKESFNAPTF		
	AWLIEWEGEKSCELGWRLVECRDINGQPASECAKEVKRPASTRPCADHPQWOL		
	GEWSKCTCGKQYKRSKLCLSHDGGVLSHESDPLKPKHFDIFCTMAECS"		
	3347..4659		
3'UTR	/gene="ADAMTS1"		
BASE COUNT	1242 a 1083 c 1255 g 1079 t		
ORIGIN			
Query Match	22.1%	Score 629.8;	DB 9; Length 4659;
Best Local Similarity	57.1%	Pred. No. 2.9e-103;	
Matches 1306;	Conservative	0; Mismatches 912;	Indels 69; Gaps 6;
QY	175	GCATTTCAGGAGACTTTTACCTACCTGACGCGCGGATGCTCAGTCTTGCGCTCCCGCC	234
Db	680	GCGTTTGACACAGCTGATCTGAGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGG	739
QY	235	TTCTCCACTGACATCTGGCGCTCCCGCTCCAGGGGCTCACCAGGGGCTC---TTCAGAC	291

Db 1820 GGGAAATGTTTGATGACAAGCCTCAGAAATCCATATCCAGCTCCAGGCGATCCCTGGC 1879

Qy 1318 GCCAGCTACACCTGAGCAGAGTCCGAGCTGGCTTTTGGCTGGCTTCCAGCCCTGT 1377

Db 1880 ACCTCGTACGATGCCAACGGGAGTCCGCTTACATTTGGGAGGAGTCCAAACACTGC 1939

Qy 1378 CTTTACATGACGTAC---TGCACAAAGCTGTGTGCACCGGAAGGCCAAGGGACAGATG 1434

Db 1940 CCCGATCCAGCCAGACATGTAGCACCTTGTGTGTACCGGCACCTCTATGTGGGTGTG 1999

Qy 1435 GTGTGCCAGACCCGCGACATTCCTCGCTGGCGGATGGACACAGCTGTGGCAAGGCAAGCTC 1494

Db 2000 GTGTGTCAAAACAAACACTTCCCGTGGGGGATGGACACAGCTGTGGAGANGGAATGG 2059

Qy 1495 TGCTCAAGGGGCTGCTGGGAGAGACACAACTCAACAAG-----CACAGGTGGAT 1548

Db 2060 TGTATCAACGGCAAGTGTGTGAACAAACCGACAGAAAGCATTTTGTATACGCTTTTCAT 2119

Qy 1549 GGTTCCTGGCCAAATGGGATCCCTATGGCCCTGTCTCGGCGACATGTGTGGGGGCGTG 1608

Db 2120 GGAAGCTGGGGAATGTGGGGGCTTGGGGAGACTGTTCCAGAACTGCGGTGGAGGATC 2179

Qy 1609 CAGCTGCCAGGAGGAGTGCACAAACCCACCCCTGCCAACGGGGCAAGTACTCGGAG 1668

Db 2180 CAGTACAGATGAGGAATGTGACAAACCCAGTCCCAAGAAATGGAGGGAAGTACTGTGA 2239

Qy 1669 GGAGTGAGGCTGAATACCGATCCTGCAATCTGGAGCCCTGCCAGCTAGCCTCCGGA 1728

Db 2240 GGCAACGAGTGGCTACAGATCCTGTACCTTGAGGACTGTCCAGAC---AATAATGGA 2296

Qy 1729 AAGAGCTTCGGGAGGAGTGTAGGCTTTCAAGGCTTACAAACAGCAGCACCACCGG 1788

Db 2297 AAAACCTTTAGAGAGGAACAATGTGAGCACACAACAGAGTTTCAAAAGCTTCTTTGGG 2356

Qy 1789 CTCACTCTCGCGTGGATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTGC 1848

Db 2357 AGTGGGCTCGGGTGGAAATGGATCCCAAGTACGCTGGCGCTCACCAAAAGGACAGTGC 2416

Qy 1849 AAGCTCATCTGCCAGGCAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGGTG 1908

Db 2417 AAGCTCATCTGCCAAGCAAAAGGCAATGGCTACTTCTCTCTTTTGCAGGCCAAAGTTGA 2476

Qy 1909 GACGGCAGCTGTGCTCTCTGACTCCACTCCGCTGTGTCTCCAGGCAAGTGCATCAAG 1968

Db 2477 GATGGTACTCCATGTAGCCCCAGATTCACCTCTGTCTGTGCAAGGACAGTGTGTAAAA 2536

Qy 1969 GCTGGCTGTATGGGAACTCGGCTCCAAGAAGATTCGACAAAGTGTGGGTGTGTGGG 2028

Db 2537 CTTGGTTGTATCCATCATAGACTCCAAAGAGAGTTTGTATTAATGTGTGTGGGG 2596

Qy 2029 GGAGACAATAAGAGCTGCAAGAAGGTGACTGGACTCTTCACCAAGCCATGCATGGCTAC 2088

Db 2597 GGAAATGGATCTACTTGTAAAAAATATCAGGATCAGTTACTAGTGCAAAACCTGGATAT 2656

Qy 2089 AATTTGTTGGTGGCCATCCCGCAGGCGCTCAGCATCGACATCCGCCAGCGGGTTAC 2148

Db 2657 CATGATATCATCAAAATCCCAACTGGAGCCCAACCAATCGAAGTGAACACGCGAACCAG 2716

Qy 2149 AAAGGCTGTACGGGATGACAATCTGCTCTGTAAGAACAGCCAAAGCAAGTACCTG 2208

Db 2717 AGGGATCCAGGAACAATGGCAGCTTCTTGCCATCAAGCTGCTGATGCCACATATAT 2776

Qy 2209 CTCACGGGCATTTCTGGTGTCTGGCGGTGGAGGGACCTGGGTGTGAAGGGCAGTCTG 2268

Db 2777 CTTAATGTGACTACACTTTGTCCACCTTAGCAGACACATTAATGTACAAGGTGTGTGC 2836

Qy 2269 CTGGGTACAGCGGACAGGCGTGGAGAGCTCGAGGCTTCCCGGCCATCTCG 2328

Db 2837 TTGAGGTACAGGGCTCTCTGCGGCATTTGAAAGAAATTCGACGCTTAGCCCTCTCAA 2896

Qy 2329 GAGCGCTGACCGGTGGAGTCTCTCGGTGGGGAAGATGACACCGCCCGCGCTCCGCTAC 2388

Db 2897 GAGCCCTTGACCATCCAGGTTCTTACTGTGGGCAATGCCCTTCGACCTAAAAATAATAC 2956

Oy 2389 TCCTTCT 2395

Db 2957 ACCTACT 2963

Search completed: May 15, 2003, 22:28:06

Job time : 9377 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 00:28:42 ; Search time 7221 Seconds
(without alignments)
11498.450 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgcttctgtggtgcatcct.....gcgtcctgaggcgtgtgctga 2853

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 24

Total number of hits satisfying chosen parameters: 14

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2802	98.2	2853	9 HSA315733	AJ315733 Homo sapi
2	2196	77.0	2930	6 AX342635	AX342635 Sequence
3	2196	77.0	2937	6 AX319860	AX319860 Sequence
4	958	33.6	157963	9 AP002986	AP002986 Homo sapi
5	958	33.6	172905	2 AC025130	AC025130 Homo sapi
6	803	28.1	170682	2 AC033429	AC033429 Homo sapi
7	724	25.4	28000	9 AP003459	AP003459 Homo sapi
8	358	12.5	172905	2 AC025130	AC025130 Homo sapi
9	288	10.1	170682	2 AC033429	AC033429 Homo sapi
10	74	2.6	182656	2 AC101990	AC101990 Mus muscu
11	61	2.1	178764	2 AC126507	AC126507 Rattus no
12	44	1.5	178764	2 AC126507	AC126507 Rattus no
13	36	1.3	182656	2 AC101990	AC101990 Mus muscu
14	35	1.2	3927	10 BC009667	BC009667 Mus muscu

ALIGNMENTS

RESULT 1
HSA315733
LOCUS Homo sapiens mRNA for metalloprotease disintegrin 15 (ADAMTS15 gene)
DEFINITION HSA315733 2853 bp mRNA linear PRI 01-MAR-2002
ACCESSION AJ315733
VERSION AJ315733.1 GI:19171175
KEYWORDS ADAMTS15 gene; disintegrin; metalloprotease; thrombospondin.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cal, S., Obaya, A.J., Llamazares, M., Garabaya, C., Quesada, V. and Lopez-Otin, C.
TITLE Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains
JOURNAL Gene 283 (1-2), 49-62 (2002)
MEDLINE 21856482
PUBMED 11867212
REFERENCE 2 (bases 1 to 2853)
AUTHORS Cal, S.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2001) Cal S., Biochemistry and Molecular Biology, University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN
FEATURES
source 1. .2853
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .2853
/gene="ADAMTS15"
1. .2853
/gene="ADAMTS15"
/codon_start=1
/product="metalloprotease disintegrin 15 with thrombospondin domains"
/protein_id="CAC86014.1"
/db_xref="GI:19171175"
/translation="MLLLGILTLAFAGRTAGGSEPEVWVPIRLIOPDINGRRYVWRG PEDSGDGLIFQITAFQEDFYHLPTDQAFAPAFSTFELGVPLAGLTGSGSDLRRCF YSGDYNAPDSEAFAYSLCGGLRGAFYGAIVYISPLNAPAPAAQRNSQAHLLQRR GVPDGSQDPTSCRGVAGSNPAILRLADPYKPRRAGFSGRRSRKRRKRFVSIPIR YVETLVADESMTKFGADLEHLTLTATARYRHPSILNPIVVKVLLLRDR SGPKTGNALTLTRNFKCAWKLNKVSQKHPEYMDTALITFRQDLCGATTCDTGMD VGTMCDPKRSCSVIEDDGLPSAFTTABELHGVNMPHONVACEVFCGLRANHMSP

TLIQIDRANPWSACSAAIITDFLDSHGDCLLDQPSKPISLPEDIPLGASYTUSQOCEL
AFVSGKPCPYMOYTKLMCTKAKGQVMYQCTRHFPWADTSCGEGKLCCLKGACVYERH
NLNKHVDSWAKWDYGPCSRTCTGGVQLARRQCTNPTPANGKYCEGVRYKRSCH
LEPCPSASAKSPRECEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKKCLICRANG
TGYYVYLAPKVDGTLTCSPDSTSVQVQKCIKAGCDNLGSKRRDKCGVCGGDNKSC
KVTYGLTFPMHGYNFVAIPAGASSIDIRQRYKGLIGDDNYLALKNSOGKYLNGH
FVVSVERDLVWKSLLRTSGTGTAVESLQASRPILLEPTVEVLSVGMATPPRVYISF
YLPKPREDKRSHPKDPRGPSVLHNSVLENOPEOPDDRPVWVAGSWGPSASCG
SGLOKRAVDCRSAGQRTVPACDAHRPVETOACGEPCTWELSAWSPCKSKCGRGFO
RRSLKCVGHGRLLRDQCNLHRKKPDELFCVLRPC"
BASE COUNT 522 a 921 c 903 g 507 t
ORIGIN

Query Match 98.2%; Score 2802; DB 9; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCATCTTAACCTCGCTTTTCGCGGGCGAACCCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGCATCTTAACCTCGCTTTTCGCGGGCGAACCCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGTGTGTTCCATCCGACTGGACCCGGACATTAACGSCCCGCCCTAC 120
DB 61 CCAGAGCGGGAGTGTGTTCCATCCGACTGGACCCGGACATTAACGSCCCGCCCTAC 120
QY 121 TACTGGCGGGGTCCCGAGGACTCGGGGGATCAGGGACTCAITTTTCAGATCAGACATTT 180
DB 121 TACTGGCGGGGTCCCGAGGACTCGGGGGATCAGGGACTCAITTTTCAGATCAGACATTT 180
QY 181 CAGGAGACTTTTACCTACACCTGACGCCGATGCTCAGTCTTTGGCTCCCGCTTCTCC 240
DB 181 CAGGAGACTTTTACCTACACCTGACGCCGATGCTCAGTCTTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATGTGGCGTCCCGCTTCCAGGGCTCACCGGGGCTCTTACAGACTGCGACGC 300
DB 241 ACTGAGCATGTGGCGTCCCGCTTCCAGGGCTCACCGGGGCTCTTACAGACTGCGACGC 300
QY 301 TGGTCTATCTGGGACGTGAACCGCGAGCGGACTCGTTGCGTGTGAGCCCTGTGC 360
DB 301 TGGTCTATCTGGGACGTGAACCGCGAGCGGACTCGTTGCGTGTGAGCCCTGTGC 360
QY 361 GGGGGGCTCCCGGAGCTTTTGGTACCGAGGGCCGAGTATGTCAATTAGCCCGCTGCC 420
DB 361 GGGGGGCTCCCGGAGCTTTTGGTACCGAGGGCCGAGTATGTCAATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGCGGCGAGCAACAGCGAGGGCCACACCTTCTCCAGCGCGG 480
DB 421 AATGCTAGCGCGCGGGCGGCGAGCAACAGCGAGGGCCACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGCGGGCTTCCGGAGACCCACACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
DB 481 GGTGTTCCGGCGGGCTTCCGGAGACCCACACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCCATCTTACGGGCGCTTGACCCCTTACAAGCGCGGGCGGCGGCTTCGGGGAG 600
DB 541 AACCCCGCCATCTTACGGGCGCTTGAGCCCTTACAAGCGCGGGCGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGAGGCTTGGCGCGCCAAAGGTTTCGTCTATATCCCGCGGTACGTG 660
DB 601 AGTCGTAGCGCGCGAGGCTTGGCGCGCCAAAGGTTTCGTCTATATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGCGGGAGCTCAATGTTCAAGTTCACGGCGGGACCTTGGAAACAT 720
DB 661 GAGACGCTGTGTGCGGGAGCTCAATGTTCAAGTTCACGGCGGGACCTTGGAAACAT 720
QY 721 TATCTGTGACGTGTGCGAACGGCGGCGACTCTTACCGCCATCCAGCATCTCTCAAC 780
DB 721 TATCTGTGACGTGTGCGAACGGCGGCGACTCTTACCGCCATCCAGCATCTCTCAAC 780
QY 781 CCCATCAACATCGTGTGTTCAAGGTGCTCTTCTTAGAGATCGTGNATCCGGGCCAAG 840
DB 781 CCCATCAACATCGTGTGTTCAAGGTGCTCTTCTTAGAGATCGTGNATCCGGGCCAAG 840

QY 841 GTACCCGGGAATGCGGCCCTGACGCTGCGCAACTTCTCTGCTGCGCAGAAAGCTGAAC 900
DB 841 GTACCCGGGAATGCGGCCCTGACGCTGCGCAACTTCTCTGCTGCGCAGAAAGCTGAAC 900
QY 901 AAAGTAGTGACAACACCCCGAGTACTGGGACACTGCACTCTCTTCCACGAGCAACAC 960
DB 901 AAAGTAGTGACAACACCCCGAGTACTGGGACACTGCACTCTCTTCCACGAGCAACAC 960
QY 961 CTGTGTGGAGCCACCACTGTGTGACACCCCTGGCATGGCTGTGTGGTACCATGTGTGAC 1020
DB 961 CTGTGTGGAGCCACCACTGTGTGACACCCCTGGCATGGCTGTGTGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGCTGTCTGTCTATTTAGGACAGATGGGCTTCCATCAGCCTTCCACACTGCC 1080
DB 1021 CCCAAGAGAAGCTGTCTGTCTATTTAGGACAGATGGGCTTCCATCAGCCTTCCACACTGCC 1080
QY 1081 CACGAGCTGGGCGACGTGTCAACATGCCCATGACATGTGAAAGTCTGTGAGGAGTG 1140
DB 1081 CACGAGCTGGGCGACGTGTCAACATGCCCATGACATGTGAAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCCCTCATCCAGATCGAOCGTGCC 1200
DB 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCCCTCATCCAGATCGAOCGTGCC 1200
QY 1201 AACCCCTGGTACAGCTGCGAGTCTGCCATCATCACGACTTCCCTGGACAGCGGCAAGGT 1260
DB 1201 AACCCCTGGTACAGCTGCGAGTCTGCCATCATCACGACTTCCCTGGACAGCGGCAAGGT 1260
QY 1261 GACTGCTCTCTGGACAAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGTGCC 1320
DB 1261 GACTGCTCTCTGGACAAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGTGCC 1320
QY 1321 AGCTACACCTGAGCGAGCTGCGAGCTGCTTTTGGCGTGGGCTCCAAAGCCCTGTGCT 1380
DB 1321 AGCTACACCTGAGCGAGCTGCGAGCTGCTTTTGGCGTGGGCTCCAAAGCCCTGTGCT 1380
QY 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGCCAAAGGACAGATGTTGTGC 1440
DB 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGCCAAAGGACAGATGTTGTGC 1440
QY 1441 CAGACCCGCGCACTTCCCTGGGCGCATGCGACCAAGCTGTGGCGAGGGCAAGCTCTGCC 1500
DB 1441 CAGACCCGCGCACTTCCCTGGGCGCATGCGACCAAGCTGTGGCGAGGGCAAGCTCTGCC 1500
QY 1501 AAAGGGGCTGGTGGAGAGACAAACCTCAACAGCAGGGGTGGATGGTTCCTGGGCG 1560
DB 1501 AAAGGGGCTGGTGGAGAGACAAACCTCAACAGCAGGGGTGGATGGTTCCTGGGCG 1560
QY 1561 AAATGGGATCCCTATGCGCCCTGCTCGCGCACATGTGTGGGGCGTGCAGCTGGCCAGG 1620
DB 1561 AAATGGGATCCCTATGCGCCCTGCTCGCGCACATGTGTGGGGCGTGCAGCTGGCCAGG 1620
QY 1621 AGGAGTGGACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
DB 1621 AGGAGTGGACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
QY 1681 AAATACCGATCTGCAATCTGAGCCCTGCCCGAGCTCAGCTCGGAAAGAGCTTCGG 1740
DB 1681 AAATACCGATCTGCAATCTGAGCCCTGCCCGAGCTCAGCTCGGAAAGAGCTTCGG 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAACCGGCTCAGCTTCGCC 1800
DB 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAACCGGCTCAGCTTCGCC 1800
QY 1801 GTGGCATGGTGGCCAACTACTTCCGGCGTGTCTCCCGGGGACAAAGTGAAGCTCATGTC 1860
DB 1801 GTGGCATGGTGGCCAACTACTTCCGGCGTGTCTCCCGGGGACAAAGTGAAGCTCATGTC 1860
QY 1861 CGAGCAATGGCAGCTGCTTCTATGCTGGCACCACCAAGGTGGTGGACGCGACGCTG 1920
DB 1861 CGAGCAATGGCAGCTGCTTCTATGCTGGCACCACCAAGGTGGTGGACGCGACGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCCGCTGTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGAT 1980

Db 1921 TGCTCTCTCTGACTCCACCTCCGCTGTGTGTCACCAAGCAAGTGCAACAAGCTGGCTGTGAT 1980
QY 1981 GGAACACCTGGGCTCCCAAGAGATTCACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
Db 1981 GGAACACCTGGGCTCCCAAGAGATTCACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGGTGACTGGACTCTTCCACCAAGCCCATGATGCTACAAATTCGTGGTG 2100
Db 2041 AGCTGCAAGAAGGTGACTGGACTCTTCCACCAAGCCCATGATGCTACAAATTCGTGGTG 2100
QY 2101 GCCATCCCGCAGGGCCCTCAAGCATGACATCCGCGCAGCCGCTTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGCAGGGCCCTCAAGCATGACATCCGCGCAGCCGCTTACAAAGGCTGATC 2160
QY 2161 GGGGATGACAACTACTCGGCTCTGAAGAACAGCCAAAGCAAGTACCTGCTCAAGGGCAT 2220
Db 2161 GGGGATGACAACTACTCGGCTCTGAAGAACAGCCAAAGCAAGTACCTGCTCAAGGGCAT 2220
QY 2221 TTGCTGTGTGTCGGGGTGGAGCGGACTGCTGCTGAAGGCACTGCTGCTCGGTACAGC 2280
Db 2221 TTGCTGTGTGTCGGGGTGGAGCGGACTGCTGCTGAAGGCACTGCTGCTCGGTACAGC 2280
QY 2281 GGCAGGGCAGCGGTGGAGAGCCTGAGGCTTCCGCGGCCATTCCTGGAGCCGCTGACC 2340
Db 2281 GGCAGGGCAGCGGTGGAGAGCCTGAGGCTTCCGCGGCCATTCCTGGAGCCGCTGACC 2340
QY 2341 GTGAGGTCTCTCTCGGTGGGAAGATGACACCGCCCGGCTGCTGCTACTCTCTATCTG 2400
Db 2341 GTGAGGTCTCTCTCGGTGGGAAGATGACACCGCCCGGCTGCTGCTACTCTCTATCTG 2400
QY 2401 CCCAAGAGCTCGGGAGGACAACTCTCTATCCCAAGGACCTCCGCGGACCCCTCTGTC 2460
Db 2401 CCCAAGAGCTCGGGAGGACAACTCTCTATCCCAAGGACCCCGGGACCCCTCTGTC 2460
QY 2461 TTGCACAAAGCGCTCAGCTCTCCAAAGAGTGGAGAGCGGAGCAGGAGCCGCT 2520
Db 2461 TTGCACAAAGCGCTCAGCTCTCCAAAGAGTGGAGAGCGGAGCAGGAGCCGCT 2520
QY 2521 GCACGCTGGGTGGTGGAGCTGGGGGCGTGTCTCCGCGAGCTGCGGCACTGGCTGCGAG 2580
Db 2521 GCACGCTGGGTGGTGGAGCTGGGGGCGTGTCTCCGCGAGCTGCGGCACTGGCTGCGAG 2580
QY 2581 AAGCGGGGTGGAGCTGCGGGGTGCGCGGGGAGCGGACGCTGCTGCTGATGCA 2640
Db 2581 AAGCGGGGTGGAGCTGCGGGGTGCGCGGGGAGCGGACGCTGCTGCTGATGCA 2640
QY 2641 GCCATCGGCCGTGGAGACACAAAGCTGCGGGGAGCGCTGCGGGGAGCGCTGCGGAGCTCAGC 2700
Db 2641 GCCATCGGCCGTGGAGACACAAAGCTGCGGGGAGCGCTGCGGGGAGCGCTGCGGAGCTCAGC 2700
QY 2701 GCCTGTACCCCTGCTCCAAAGAGCTGCGGGGCGGGGATTTTCAGAGGGCTCACTCAAGTGT 2760
Db 2701 GCCTGTACCCCTGCTCCAAAGAGCTGCGGGGCGGGGATTTTCAGAGGGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGAGGCGGGCTGCTGCGGGGAGCGCTGCGGGGAGCGCTGCGGGGAGCGGAG 2820
Db 2761 GTGGGCGAGGCGGGCTGCTGCGGGGAGCGCTGCGGGGAGCGCTGCGGGGAGCGGAG 2820
QY 2821 GAGCTGACTTCTGCTGCTGAGCGGCTGCTGA 2853
Db 2821 GAGCTGACTTCTGCTGCTGAGCGGCTGCTGA 2853

RESULT 2
AX342635
LOCUS
DEFINITION
Sequence 32 from Patent WO0198468.
ACCESSION
AX342635
VERSION
AX342635.1
KEYWORDS
GI:18152032
SOURCE
human.
ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Yue, H., Elliott, V. S., Gandhi, A. R., Lal, P., Au-Young, J.,
Tribouley, C. M., Deleage, A. M., Baughn, M. R., Nguyen, D. B., Lee, E. A.,
Hafalia, A., Khan, F. A., Walla, N. K., Yao, M. G., Lu, D. A., Patterson, C.,
Tang, Y. T., Walsh, R. T., Azimzai, J., Ramkumar, J., Xu, Y., and Reddy, R.
Patent: WO 0198468-A 32 27-DEC-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..2930
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7473089CB1"
BASE COUNT 529 a 958 c 931 g 512 t
ORIGIN
Query Match 77.0%; Score 2196; DB 6; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 ATGCTTCTCTGGGCACTCTAACCCTTGGCTTTGCGCGGGGAAACCGCTGGAGTCTGTAG 60
Db 75 ATGCTTCTCTGGGCACTCTAACCCTTGGCTTTGCGCGGGGAAACCGCTGGAGTCTGTAG 134
QY 61 CCAGAGCGGGAGGTAGTCTGCTCCATCCGACCTGGACCGGAGATTAAAGCGCGCGCTAC 120
Db 135 CCAGAGCGGGAGGTAGTCTGCTCCATCCGACCTGGACCGGAGATTAAAGCGCGCGCTAC 194
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGACTCATTTTTCAGATCACAGATTAT 180
Db 195 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGACTCATTTTTCAGATCACAGATTAT 254
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTAGTCTTGGCTCCCGCTTTC 240
Db 255 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTAGTCTTGGCTCCCGCTTTC 314
QY 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTGAGCGC 300
Db 315 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTGAGCGC 374
QY 301 TGTCTTATTTTGGGAGCTGAACCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 375 TGTCTTATTTTGGGAGCTGAACCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
QY 361 GGGGGCTCCCGGAGCTTTGGCTACGAGCGCGGAGTATGCTATGAGCGCTGCTGCTGCT 420
Db 435 GGGGGCTCCCGGAGCTTTGGCTACGAGCGCGGAGTATGCTATGAGCGCTGCTGCTGCT 494
QY 421 AATGCTAGCGCGCGGCGGAGCTGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 495 AATGCTAGCGCGCGGCGGAGCTGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
QY 481 GGTGTTCGGGGGCGCTTCCGAGAGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 555 GGTGTTCGGGGGCGCTTCCGAGAGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
QY 541 AACCCCGCATCTACGCGGCTTACAGCGCGGCGGCGGCGGCTTCCGGGAG 600
Db 615 AACCCCGCATCTACGCGGCTTACAGCGCGGCGGCGGCGGCTTCCGGGAG 674
QY 601 AGTCTAGCGCGGCGGAGCTTGGGCGCGGAGCTTTCGCTGCTATCCCGCGGTACGCTG 660
Db 675 AGTCTAGCGCGGCGGAGCTTGGGCGCGGAGCTTTCGCTGCTATCCCGCGGTACGCTG 734
QY 661 GAGAGCTGTGTGCTGCGGAGCTCAATGTCAAGTTCACGCGCGGAGCTGGAACAT 720
Db 735 GAGAGCTGTGTGCTGCGGAGCTCAATGTCAAGTTCACGCGCGGAGCTGGAACAT 794
QY 721 TATCTGTGAGCTGTGCGCAACGCGGCGGAGCTTACCGCCATCCAGCATCTCTAAC 780
Db 795 TATCTGTGAGCTGTGCGCAACGCGGCGGAGCTTACCGCCATCCAGCATCTCTAAC 854

```
QY 781 CCCATCAACATCGTTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCGGGCCCAAG 840
Db 855 CCCATCAACATCGTTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCGGGCCCAAG 914
QY 841 GTCACCGCAATCGGGCCCTGACCTGGCAACTTCTGCTGGCAGAGAAAGCTCAAC 900
Db 915 GTCACCGCAATCGGGCCCTGACCTGGCAACTTCTGCTGGCAGAGAAAGCTCAAC 974
QY 901 AAAGTGAGTGACAAAGCAACCCCGAGTACTGGGACACTGCCATCTCTTCAACGACGAGAC 960
Db 975 AAAGTGAGTGACAAAGCAACCCCGAGTACTGGGACACTGCCATCTCTTCAACGACGAGAC 1034
QY 961 CTGTGTGAGGACCAACCTGTGACACCTGGGCAATGCTGATGGGTACCATGTGTGAC 1020
Db 1035 CTGTGTGAGGACCAACCTGTGACACCTGGGCAATGCTGATGGGTACCATGTGTGAC 1094
QY 1021 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGTTCATCAGGCTTCAACCACTGCC 1080
Db 1095 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGTTCATCAGGCTTCAACCACTGCC 1154
QY 1081 CACGAGCTGGGCCAGTGTTCACATGCCCATGACCAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1155 CACGAGCTGGGCCAGTGTTCACATGCCCATGACCAATGTGAAGTCTGTGAGGAGGTG 1214
QY 1141 TTGTGGAAAGCTCCGAGCAACCAACATGATGCCCGACCTCATCCAGATCGACCGTGC 1200
Db 1215 TTGTGGAAAGCTCCGAGCAACCAACATGATGCCCGACCTCATCCAGATCGACCGTGC 1274
QY 1201 AACCCCTGCTGAGCTGCAGTGCCTCATCATCACCGACTTCTTGGACAGCGGACAGGT 1260
Db 1275 AACCCCTGCTGAGCTGCAGTGCCTCATCATCACCGACTTCTTGGACAGCGGACAGGT 1334
QY 1261 GACTGCCCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGGGCC 1320
Db 1335 GACTGCCCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGGGCC 1394
QY 1321 AGCTACACCTTGAGCCAGCTGCGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCTCT 1380
Db 1395 AGCTACACCTTGAGCCAGCTGCGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCTCT 1454
QY 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1440
Db 1455 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1514
QY 1441 CAGACCCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGCGGAGGCAAGCTCTGCTC 1500
Db 1515 CAGACCCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGCGGAGGCAAGCTCTGCTC 1574
QY 1501 AAAGGGGCTGCTGGAGAGACAAACCTCAACAAGCACAGGCTGGATGTTCTTGGGCC 1560
Db 1575 AAAGGGGCTGCTGGAGAGACAAACCTCAACAAGCACAGGCTGGATGTTCTTGGGCC 1634
QY 1561 AAATGGGATCCCTATFGCCCTGTGTCGCGACATGTGTGGGGGCGTGCAGCTGGCCAGG 1620
Db 1635 AAATGGGATCCCTATFGCCCTGTGTCGCGACATGTGTGGGGGCGTGCAGCTGGCCAGG 1694
QY 1621 AGGCAGTGCACCAACCCCACTCCCAACCGGGGCAAGTACTCCGAGGGAGTGAAGGTG 1680
Db 1695 AGGCAGTGCACCAACCCCACTCCCAACCGGGGCAAGTACTCCGAGGGAGTGAAGGTG 1754
QY 1681 AAATACCGATCTGCAATCTTGAGCCCTGCCCCAGCTCAGGCTTCGGAAAGAGCTTCCGG 1740
Db 1755 AAATACCGATCTGCAATCTTGAGCCCTGCCCCAGCTCAGGCTTCGGAAAGAGCTTCCGG 1814
QY 1741 GAGGAGCAGTGTAGGCTTTCAAGGGCTTACAAACCAACAGCAACCCGGCTCACTCTGCC 1800
Db 1815 GAGGAGCAGTGTAGGCTTTCAAGGGCTTACAAACCAACAGCAACCCGGCTCACTCTGCC 1874
QY 1801 GTGGCATGGTCCCAAGTACTCCGGCTGTCTCCCGGGGACAGTGCAGCTCATCTGC 1860
Db 1875 GTGGCATGGTCCCAAGTACTCCGGCTGTCTCCCGGGGACAGTGCAGCTCATCTGC 1934
QY 1861 CGAGCAATGGCACTGCTACTTCTATGTGCTGGCACCCCAA ---GGTGGTGAGCGGACG 1917
```

```
Db 1935 CGAGCAATGGCACTGGCTACTTCTATGTGTCGCCCAAGGTGGTGGTGGACGCTACG 1994
QY 1918 CTGTGCTCTCCTGACTCCACCTCCGCTGTGTCCAAAGGCAAGTGCATCAAGGCTGCTGT 1977
Db 1995 CTGTGCTCTCCTGACTCCACCTCCGCTGTGTCCAAAGGCAAGTGCATCAAGGCTGCTGT 2054
QY 1978 GATGGAACTGGGCTCCCAAGAGAGATTCGACAAGTGTGGGGTGTGGGGGAGACAAT 2037
Db 2055 GATGGAACTGGGCTCCCAAGAGAGATTCGACAAGTGTGGGGTGTGGGGGAGACAAT 2114
QY 2038 AAGAGTGAAGAAGTGTGACTGACTTTCACCAAGCCCATGCATGGCTACAAATTTCTG 2097
Db 2115 AAGAGTGAAGAAGTGTGACTGACTTTCACCAAGCCCATGCATGGCTACAAATTTCTG 2174
QY 2098 GTGGCCATCCCGCAGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGGCTG 2157
Db 2175 GTGGCCATCCCGCAGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGGCTG 2234
QY 2158 ATCGGGGATGACAATACCTGCTCTGAAGAACAGCCAGGCAAGTACCTGCTCAACGG 2217
Db 2235 ATCGGGGATGACAATACCTGCTCTGAAGAACAGCCAGGCAAGTACCTGCTCAACGG 2294
QY 2218 CATTTCTGTGTGCGCGGTGGAGCGGACCTGTGTGTAAGGGCAGTCTGTGCGGTAC 2277
Db 2295 CATTTCTGTGTGCGCGGTGGAGCGGACCTGTGTGTAAGGGCAGTCTGTGCGGTAC 2354
QY 2278 AGCGGCACGGGCACAGCGGTGAGAGCCTGAGGCTTCCCGGCCCATCTTGAGGCGCTG 2337
Db 2355 AGCGGCACGGGCACAGCGGTGAGAGCCTGAGGCTTCCCGGCCCATCTTGAGGCGCTG 2414
QY 2338 ACCGTGGAGTCTCTCCGTGGGAGAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2397
Db 2415 ACCGTGGAGTCTCTCCGTGGGAGAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2474
QY 2398 CTGCCCCAAAGAGCTCGGGAGGACAACTCTCTCATCC 2436
Db 2475 CTGCCCCAAAGAGCTCGGGAGGACAACTCTCTCATCC 2513

RESULT 3
AX319860
LOCUS
DEFINITION Sequence 24 from Patent WO0183782.
ACCESSION AX319860
VERSION AX319860.1 GI:17901450
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Plowman, G.D., Whyte, D., Sudarsanam, S., Manning, G., Caenepeel, S. and
Payne, V.
Novel proteases
Patent: WO 0183782-A 24 08-NOV-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
1..2937
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 531 a 958 c 936 g 512 t
ORIGIN

Query Match 77.0%; Score 2196; DB 6; Length 2937;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 ATGCTTCTGCTGGGATCTTAACCTTGGCTTTCGCGGGGCAACCGCTGAGGCTCTGAG 60
Db 76 ATGCTTCTGCTGGGATCTTAACCTTGGCTTTCGCGGGGCAACCGCTGAGGCTCTGAG 135
QY 61 CCAGAGCGGAGGAGTAGTCTCCATCCGATCCGATGGACCCCGGACATTAACGGCCGCCCTAC 120
```

|||||
Db 136 CCAGAGCGGAGGTAGTCGTTCCCATCCGACTGGAGCCGGACATTAAAGCGCGCGCTAC 195
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCACAGCATTT 180
Db 196 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCACAGCATTT 255
QY 181 CAGGAGGACTTTTACTACATCAGTACGCGGATGCTCAGTTCTTTGGTCCGCGCTTCGC 240
Db 256 CAGGAGGACTTTTACTACATCAGTACGCGGATGCTCAGTTCTTTGGTCCGCGCTTCGC 315
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGTCAACGGGGCTTCAGAGCTGGGACGC 300
Db 316 ACTGAGCATCTGGGCGTCCCTCCAGGGGTCAACGGGGCTTCAGAGCTGGGACGC 375
QY 301 TGTCTTATTTCTGGGAGCTGAACCCGAGCGGACTCTGTTGCTGTGCTGTCGAGCTATGC 360
Db 376 TGTCTTATTTCTGGGAGCTGAACCCGAGCGGACTCTGTTGCTGTGCTGTCGAGCTATGC 435
QY 361 GGGGGCTCCGGAGCTTTGGCTACCGAGCGCGCGAGTATGTCTATTAGCTGCTGCC 420
Db 436 GGGGGCTCCGGAGCTTTGGCTACCGAGCGCGCGAGTATGTCTATTAGCTGCTGCC 495
QY 421 ATGCTAGCGCGCGCGCGCGAGCGCAACAGCCAGGCGCACACTTCTCCAGCGCGG 480
Db 496 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCCAGGCGCACACTTCTCCAGCGCGG 555
QY 481 GGTGTTCCGGGGGGGCTTCCGGAGACCCAGCTCTCGCTCGGGGGTGAGCTCGGGGTG 540
Db 556 GGTGTTCCGGGGGGGCTTCCGGAGACCCAGCTCTCGCTCGGGGGTGAGCTCGGGGTG 615
QY 541 AACCCCGCATCTCTAGGCGCTTGACACCTTACAGCGCGCGGGGGCTTCGGGGAG 600
Db 616 AACCCCGCATCTCTAGGCGCTTGACACCTTACAGCGCGCGGGGGCTTCGGGGAG 675
QY 601 AGTCGTAGCGCGCGAGTCTGGGCGCGCAAGCGTTTCGTGTCTATCCCGCGTACGTG 660
Db 676 AGTCGTAGCGCGCGAGTCTGGGCGCGCAAGCGTTTCGTGTCTATCCCGCGTACGTG 735
QY 661 GAGACGCTGGTGTGCGGACAGTCAATGTCAGTTCCAGCGCGGAGCTTGAACAT 720
Db 736 GAGACGCTGGTGTGCGGACAGTCAATGTCAGTTCCAGCGCGGAGCTTGAACAT 795
QY 721 TATCTGCTGACGCTGCTGCAACGGCGCGGACTCTACCGCCATCCAGATCCTCAAC 780
Db 796 TATCTGCTGACGCTGCTGCAACGGCGCGGACTCTACCGCCATCCAGATCCTCAAC 855
QY 781 CCCATCAACATGTTGTGTCAGTGTGCTTCTTAGAGATGCTGACTCCGGCGCAAG 840
Db 856 CCCATCAACATGTTGTGTCAGTGTGCTTCTTAGAGATGCTGACTCCGGCGCAAG 915
QY 841 GTCACCGCAATCGCGCCCTGACGCTGCGCAACTTCTGTTCGTGCGAGAAAGCTGAAC 900
Db 916 GTCACCGCAATCGCGCCCTGACGCTGCGCAACTTCTGTTCGTGCGAGAAAGCTGAAC 975
QY 901 AAGTGCAGTACAGCAACCCAGTACTGGGACTGCCATCTCTTCAACAGCAGGAC 960
Db 976 AAGTGCAGTACAGCAACCCAGTACTGGGACTGCCATCTCTTCAACAGCAGGAC 1035
QY 961 CTGTGTGGAGCCACCACTGTGACACCTCTGGCATGGGTGATGAGGTACATGTGTGAC 1020
Db 1036 CTGTGTGGAGCCACCACTGTGACACCTCTGGCATGGGTGATGAGGTACATGTGTGAC 1095
QY 1021 CCCAAGAGAGTGTCTGTCTATTGAGGAGGATGGGCTTCCATCAGCCTTACACATGCC 1080
Db 1096 CCCAAGAGAGTGTCTGTCTATTGAGGAGGATGGGCTTCCATCAGCCTTACACATGCC 1155
QY 1081 CAGGAGTGGGCGACGTTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGGTG 1140
Db 1156 CAGGAGTGGGCGACGTTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGGTG 1215
QY 1141 TTTGGAGAGTCCGAGCCCAACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGCC 1200
|||||

Db 1216 TTTGGAAAGCTCCGAGCAACCATCATGATGTCCCGACCCCTCATCCAGATCGACCGTGC 1275
QY 1201 AACCCCTGTCAGCTGAGCTGCAGTGTCCCATCATCACCGACTTCTCTGGACAGCGGACGGT 1260
Db 1276 AACCCCTGTCAGCTGAGCTGCAGTGTCCCATCATCACCGACTTCTCTGGACAGCGGACGGT 1335
QY 1261 GACTCCCTCTGGAGCAACACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
Db 1336 GACTCCCTCTGGAGCAACACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1395
QY 1321 AGCTACACCTGAGCAGCAGTGGAGCTGCTTTTGGCTGGGCTCCAAAGCCCTGTCCT 1380
Db 1396 AGCTACACCTGAGCAGCAGTGGAGCTGCTTTTGGCTGGGCTCCAAAGCCCTGTCCT 1455
QY 1381 TACATGCAGTACTGCACCAAGCTGTGTGCACCGGGAAGCCAAAGGACAGATGGTGTGC 1440
Db 1456 TACATGCAGTACTGCACCAAGCTGTGTGCACCGGGAAGCCAAAGGACAGATGGTGTGC 1515
QY 1441 CAGACCCGCCACTTCCCTGGCCGATGGCAGCAGCTGTGGCGAGGCAAGCTCTGCCCTC 1500
Db 1516 CAGACCCGCCACTTCCCTGGCCGATGGCAGCAGCTGTGGCGAGGCAAGCTCTGCCCTC 1575
QY 1501 AAAGGGGCTGCTGGAGAGACACAACCTCAACAGCACAGGGGTGTGATGTTCTTGGGCC 1560
Db 1576 AAAGGGGCTGCTGGAGAGACACAACCTCAACAGCACAGGGGTGTGATGTTCTTGGGCC 1635
QY 1561 AANTGGATCCCTATGCGCCCTGCTCGCACATGTGTGGGGGCTGCGAGCTGCCAGG 1620
Db 1636 AANTGGATCCCTATGCGCCCTGCTCGCACATGTGTGGGGGCTGCGAGCTGCCAGG 1695
QY 1621 AGGAGTGCACCAACCCGCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGAAGGTG 1680
Db 1696 AGGAGTGCACCAACCCGCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGAAGGTG 1755
QY 1681 AATACCGATCTCGAATCTGAGGCCCTGCCCGAGCTCAGCCTCCGGAAGAGTTCGG 1740
Db 1756 AATACCGATCTCGAATCTGAGGCCCTGCCCGAGCTCAGCCTCCGGAAGAGTTCGG 1815
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACACAGCACACACCAACCCGGCTCACCTCGCC 1800
Db 1816 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACACAGCACACACCAACCCGGCTCACCTCGCC 1875
QY 1801 GTGGCANTGGTCCCAAGTACTCCGGGCTGTCTCCCGGAGCAAGTGAAGCTCATCTGC 1860
Db 1876 GTGGCANTGGTCCCAAGTACTCCGGGCTGTCTCCCGGAGCAAGTGAAGCTCATCTGC 1935
QY 1861 CGAGCCAATGGCACTGGGTACTTCTATGTGTGGCACCCAA ---GGTGTGGACGCAAG 1917
Db 1936 CGAGCCAATGGCACTGGGTACTTCTATGTGTGGCACCCCAAGGTGTGTGGAGCGCAG 1995
QY 1918 CTGTGCTCTCTGACTCCACCTCCGCTGTGTGTCGAAGGCAAGTGCATCAAGCTGTGTGT 1977
Db 1996 CTGTGCTCTCTGACTCCACCTCCGCTGTGTGTCGAAGGCAAGTGCATCAAGCTGTGTGT 2055
QY 1978 GATGGAACTGGGCTTCAAGAGAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAT 2037
Db 2056 GATGGAACTGGGCTTCAAGAGAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAT 2115
QY 2038 AAGAGCTGCAAGAGTGTGACTGCTTCAACAGCCCATGTCATGGGTACAAATTCGTGT 2097
Db 2116 AAGAGCTGCAAGAGTGTGACTGCTTCAACAGCCCATGTCATGGCTACAAATTCGTGT 2175
QY 2098 GTGGCCATCCCGCAGGCGCTCAAGCATCCGACAGCGGTTTCAAGAGCGCTG 2157
Db 2176 GTGGCCATCCCGCAGGCGCTCAAGCATCCGACAGCGGTTTCAAGAGCGCTG 2235
QY 2158 ATCGGGATGACAATACCTGCTCTGAAGAACGCCAAGCAAGTACTCTCAACGGG 2217
Db 2236 ATCGGGATGACAATACCTGCTCTGAAGAACGCCAAGCAAGTACTCTCAACGGG 2295
QY 2218 CATTTGCTGTGTGGCGGTGGAGGGACCTGTGTGTGAAGGGCAGTCTCTCGGTAC 2277
Db 2296 CATTTGCTGTGTGGCGGTGGAGGGACCTGTGTGTGAAGGGCAGTCTCTCGGTAC 2355
|||||

```
QY 2278 AGCGGACGGGCACAGCGGTGGAGAGCCTGCAGGCTTCCGGGCCCATCTTGGAGCGCGTG 2337
|||||
Db 2356 AGCGGACGGGCACAGCGGTGGAGAGCCTGCAGGCTTCCGGGCCCATCTTGGAGCGCGTG 2415
|||||
QY 2338 ACCGTGAGGTCTCTCTCCGTGGGAGAGACACCGCCCGGGTCCGCTACTCTTCTAT 2397
|||||
Db 2416 ACCGTGAGGTCTCTCTCCGTGGGAGAGACACCGCCCGGGTCCGCTACTCTTCTAT 2475
|||||
QY 2398 CTGCCCCAAAGAGCTCGGGAGGACAAGTCTCTCATCCC 2436
|||||
Db 2476 CTGCCCCAAAGAGCTCGGGAGGACAAGTCTCTCATCCC 2514
|||||

RESULT 4
AP002986 157963 bp DNA linear PRI 08-MAR-2002
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-121M22,
DEFINITION complete sequence.
ACCESSION AP002986
VERSION AP002986.2 GI:19263031
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-121M22.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 157963)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Mar 7, 2002 this sequence version replaced gi:11559301.
FEATURES
source
location/Qualifiers
1..157963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-121M22"

BASE COUNT 39742 a 36722 c 37247 g 44252 t
ORIGIN

Query Match 33.6%; Score 958; DB 9; Length 157963;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
Db 134115 ATGCTTCTGCTGGGATCCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 134174

QY 61 CCAGAGCGGAGGTAGTCTGCCATCCAGCTGACCGCGGACATTAACGGCCGCGCTAC 120
Db 134175 CCAGAGCGGAGGTAGTCTGCCATCCAGCTGACCGCGGACATTAACGGCCGCGCTAC 134234

QY 121 TACTGGCGGGGTCCCGAGGACTCCGGGATCAGGAGCTCAFTTTTCAGATCACAGCATTT 180
Db 134235 TACTGGCGGGGTCCCGAGGACTCCGGGATCAGGAGCTCAFTTTTCAGATCACAGCATTT 134294

QY 181 CAGGAGGACTTTTACCTACACTGACCGCGATGCTAGTCTTGGCTCCCGCTTCTCC 240
Db 134295 CAGGAGGACTTTTACCTACACTGACCGCGATGCTAGTCTTGGCTCCCGCTTCTCC 134354

QY 241 ACTGAGCATCTGGGGTCCCGCTCCAGGGCTTCACCGGGGCTTTCAGACCTCGCAGCG 300
|||||
```

```
Db 134355 ACTGAGCATCTGGGGTCCCGCTCCAGGGGCTACCGGGGGCTCTTCAGACCTCGCAGCG 134414
QY 301 TGCCTTCTATTTGGGGACGTGAAGCCCGAGCCGAGCTCGTTTCGCTGCTGTGAGCCTGTGC 360
|||||
Db 134415 TGCCTTCTATTTGGGGACGTGAAGCCCGAGCCGAGCTCGTTTCGCTGCTGTGAGCCTGTGC 134474
|||||
QY 361 GGGGGGCTCCCGGGAGCCTTTGGGTACCGAGGCGCGCGAGTATGTCATTTAGCCCGCTGCC 420
|||||
Db 134475 GGGGGGCTCCCGGGAGCCTTTGGGTACCGAGGCGCGCGAGTATGTCATTTAGCCCGCTGCC 134534
|||||
QY 421 AATGCTAGCGCGCGCGCGCGAGCAACAGCCAGGCGCACACCTTCTCCAGCGCGGG 480
|||||
Db 134535 AATGCTAGCGCGCGCGCGCGAGCAACAGCCAGGCGCACACCTTCTCCAGCGCGGG 134594
|||||
QY 481 GGTGTTCCGGGGCGGGCTTCCGGGAGACCCACCTCTCTGCTCGGGGGTGGCTTCGGGCTGG 540
|||||
Db 134595 GGTGTTCCGGGGCGGGCTTCCGGGAGACCCACCTCTCTGCTCGGGGGTGGCTTCGGGCTGG 134654
|||||
QY 541 AACCCCGGCATCTTACGGGCGCTTGACCTTTACAGCCCGCGGGCGGGGCTTCGGGGAG 600
Db 134655 AACCCCGGCATCTTACGGGCGCTTGACCTTTACAGCCCGCGGGCGGGGCTTCGGGGAG 134714
|||||
QY 601 AGTCGTAGCGCGCGCGAGGTCTGGGCGCGCCAAAGCGTTTCTGCTTATCCCGCGGTACGTG 660
|||||
Db 134715 AGTCGTAGCGCGCGCGAGGTCTGGGCGCGCCAAAGCGTTTCTGCTTATCCCGCGGTACGTG 134774
|||||
QY 661 GAGAGCGTGTGCTGCTCGCGAGCAGTCAATGTGTCCTTCCAGCGCGCGGAGCTTGAACAT 720
Db 134775 GAGAGCGTGTGCTGCTCGCGAGCAGTCAATGTGTCCTTCCAGCGCGCGGAGCTTGAACAT 134834
|||||
QY 721 TATCTGCTGACGCTGCTGGCAACGGCGCGGAGCTTACCGCCATCCAGCATCTCTCAAC 780
Db 134835 TATCTGCTGACGCTGCTGGCAACGGCGCGGAGCTTACCGCCATCCAGCATCTCTCAAC 134894
|||||
QY 781 CCATCAACATCGTTGTGGTGAAGTGTGCTTCTTAGAGATCGTGACTCCGGGCGCCCAAG 840
Db 134895 CCATCAACATCGTTGTGGTGAAGTGTGCTTCTTAGAGATCGTGACTCCGGGCGCCCAAG 134954
|||||
...QY 841 GTCCCGGCAATCGCGCGCTGACGTGCGCAACTTCTGTGCTGCGCAGAGAAGCTGAAC 900
Db 134955 GTCCCGGCAATCGCGCGCTGACGTGCGCAACTTCTGTGCTGCGCAGAGAAGCTGAAC 135014
|||||
QY 901 AAGTGTAGTGACAAGCACCCCGAGTACTGGGACACTGCATCTCTTCCACAGGCGAG 958
Db 135015 AAGTGTAGTGACAAGCACCCCGAGTACTGGGACACTGCATCTCTTCCACAGGCGAG 135072
|||||
```

RESULT 5

```
AC025130/c 172905 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 11 clone RP11-211H6 map 11, WORKING DRAFT
DEFINITION SEQUENCE, 16 unordered pieces.
ACCESSION AC025130
VERSION AC025130.2 GI:7387384
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 172905)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-211H6
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 172905)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Bohuslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campoliano,A., Castle,A., Choehel,Y., Colangelo,M., Collins,S.,
Collumore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glend,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
```

Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrim, J., Menues, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Direct Submission
 Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 172905)

Birren, B., Linton, L., Nusbaum, C., Landers, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, P.,
 Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrim, J., Menues, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 3, 2000 this sequence version replaced gl:17158941.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7808

Center clone name: 211_H_6

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 161899 bases at least Q40

Consensus quality: 167402 bases at least Q30

Consensus quality: 169801 bases at least Q20

Insert size: 177000; agarose-fp

Insert size: 171405; sum-of-contigs

Quality coverage: 4.1 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1

21: contig of 21 bp in length

* 122 121: gap of 100 bp
 * 122 1343: contig of 1222 bp in length
 * 1344 1443: gap of 100 bp
 * 1444 2636: contig of 1193 bp in length
 * 2637 2736: gap of 100 bp
 * 2737 4052: contig of 1316 bp in length
 * 4053 4152: gap of 100 bp
 * 4153 8044: contig of 3892 bp in length
 * 8045 8144: gap of 100 bp
 * 8145 14949: contig of 6805 bp in length
 * 14950 15049: gap of 100 bp
 * 15050 23535: contig of 8486 bp in length
 * 23536 23635: gap of 100 bp
 * 23636 30261: contig of 6626 bp in length
 * 30262 30361: gap of 100 bp
 * 30362 40941: contig of 10580 bp in length
 * 40942 41041: gap of 100 bp
 * 41042 54893: contig of 13852 bp in length
 * 54894 54993: gap of 100 bp
 * 54994 68888: contig of 13895 bp in length
 * 68889 68988: gap of 100 bp
 * 68989 83702: contig of 14714 bp in length
 * 83703 83802: gap of 100 bp
 * 83803 98393: contig of 14591 bp in length
 * 98394 98493: gap of 100 bp
 * 98494 117356: contig of 18863 bp in length
 * 117357 117456: gap of 100 bp
 * 117457 140589: contig of 23133 bp in length
 * 140590 140689: gap of 100 bp
 * 140690 172905: contig of 32216 bp in length.

FEATURES

source

1..172905
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="11"
 /clone="RP11-211H6"
 /clone_lib="RPC1-11 Human Male BAC"

misc_feature

1..21
 /note="assembly_fragment"
 clone_end:17
 vector_side:right

misc_feature

122..1343
 /note="assembly_fragment"
 1444..2636

misc_feature

2737..4052
 /note="assembly_fragment"

misc_feature

4153..8044
 /note="assembly_fragment"

misc_feature

8145..14949
 /note="assembly_fragment"

misc_feature

15050..23535
 /note="assembly_fragment"

misc_feature

23636..30261
 /note="assembly_fragment"

misc_feature

30362..40941
 /note="assembly_fragment"

misc_feature

41042..54893
 /note="assembly_fragment"

misc_feature

54994..68888
 /note="assembly_fragment"

misc_feature

68989..83702
 /note="assembly_fragment"

misc_feature

83803..98393
 /note="assembly_fragment"

misc_feature

98494..117356
 /note="assembly_fragment"

misc_feature

117457..140589
 /note="assembly_fragment"

misc_feature

140690..172905

RESULT 6	AC023429	LOCUS	DEFINITION	AC023429	Homo sapiens chromosome 11 clone RP11-121M22, WORKING DRAFT	170682 bp	DNA	linear	HTG 10-APR-2001
ACCESSION	AC023429	VERSION	KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.	GI:13569974	7	unordered pieces.		
SOURCE	ORGANISM	REFERENCE	AUTHORS	Unpublished	2 (bases 1 to 170682)				
JOURNAL	AUTHORS	TITLE	JOURNAL	Submitted (14-FEB-2000)	DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
COMMENT				On Apr 10, 2001 this sequence version replaced gi:13562078.					
				Center: Stanford DNA Sequencing and Technology Development					
				Center code: SDSTDC					
				Web site: http://sequence-www.stanford.edu/group/human/					
				Contact: hum-info@sequence.stanford.edu					
				----- Project Information					
				Center project name: 837					
				Center clone name: RP11-121M22					
				----- Summary Statistics					
				Chemistry: Dye-primer; 12% of reads					
				Chemistry: Dye-terminator Big Dye; 86% of reads					
				Assembly program: Phrap; version 0.990319					
				Consensus quality: 165770 bases at least Q40					
				Consensus quality: 167255 bases at least Q30					
				Consensus quality: 167918 bases at least Q20					
				Insert size: 172423; agarose-fp					
				Insert size: 170082; sum-of-contigs					
				Quality coverage: 9.0x in Q20 bases; agarose-fp					
				Quality coverage: 9.1x in Q20 bases; sum-of-contigs					
				* NOTE: This is a 'working draft' sequence. It currently					
				* consists of 7 contigs. The true order of the pieces					
				* is not known and their order in this sequence record is					
				* arbitrary. Gaps between the contigs are represented as					
				* runs of N, but the exact sizes of the gaps are unknown.					
				* This record will be updated with the finished sequence					
				* as soon as it is available and the accession number will					
				* be preserved.					
				1 1856: contig of 1856 bp in length					
				* 1857 1956: gap of unknown length					
				* 1957 7065: contig of 5109 bp in length					
				* 7066 7165: gap of unknown length					
				* 7166 21089: contig of 13924 bp in length					
				* 21090 21189: gap of unknown length					
				* 21190 34145: contig of 12956 bp in length					
				* 34146 34245: gap of unknown length					
				* 34246 48562: contig of 14317 bp in length					
				* 48563 48662: gap of unknown length					
				* 48663 109179: contig of 60517 bp in length					
				* 109180 109279: gap of unknown length					
				* 109280 170682: contig of 61403 bp in length.					
				Location/Qualifiers					
				1. 170682					
				FEATURES					
				Source					


```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone_lib="RPII human BAC library 11"
1..1856
/feature="assembly_name:Contig25"
1957..7065
/feature="assembly_name:Contig26"
7166..21089
/feature="assembly_name:Contig27"
21190..34145
/feature="assembly_name:Contig28"
34246..48562
/feature="assembly_name:Contig29"
48663..109179
/feature="assembly_name:Contig30"
109280..170682
/feature="assembly_name:Contig31"
clone_end:T7"
BASE COUNT 43947 a 39732 c 39514 g 46884 t 605 others
ORIGIN
Query Match 28.1%; Score 803; DB 2; Length 170682;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 156 ACTCATTTTTCAGATCAGACGATTCAGGAGGACTTTTACCTACACCTGACGCGCGGATGC 215
Db 98615 ACTCATTTTTCAGATCAGACGATTCAGGAGGACTTTTACCTACACCTGACGCGCGGATGC 98674
QY 216 TCAGTTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGCGCTCCCGCTCCAGGGCTCAC 275
Db 98675 TCAGTTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGCGCTCCCGCTCCAGGGCTCAC 98734
QY 276 CGGGGGCTTTCAGACCTGCGACGCTGCTTATCTTGGGGAGCTGAACCGCGAGCGGA 335
Db 98735 CGGGGGCTTTCAGACCTGCGACGCTGCTTATCTTGGGGAGCTGAACCGCGAGCGGA 98794
QY 336 CTCGTCTGCTGTGAGCCTGTGGGGGGCTCCGGGGAGCCTTTGGCTACCGAGCGC 395
Db 98795 CTCGTCTGCTGTGAGCCTGTGGGGGGCTCCGGGGAGCCTTTGGCTACCGAGCGC 98854
QY 396 CGAGTATGCTATAGCCGCTGCCAATGCTAGCGCCCGCGCGAGCGCGAGCGCAACGCCA 455
Db 98855 CGAGTATGCTATAGCCGCTGCCAATGCTAGCGCCCGCGCGAGCGCGAGCGCAACGCCA 98914
QY 456 GGGGCGACACCTTCTCCAGCGCCGGGTGTTCGGGGGGGGCTTCCGGGAGACCCACCTC 515
Db 98915 GGGGCGACACCTTCTCCAGCGCCGGGTGTTCGGGGGGGGCTTCCGGGAGACCCACCTC 98974
QY 516 TCGTGGGGGTGGCTCGGGCTGGAACCCCGCCATCTACGGGCGCTTGGAGCCTTAA 575
Db 98975 TCGTGGGGGTGGCTCGGGCTGGAACCCCGCCATCTACGGGCGCTTGGAGCCTTAA 99034
QY 576 CCCGGGGGGGGCTTGGGGAGAGTCTGTAGCGCGCGAGGCTGGGGCGCGCAAGCG 635
Db 99035 CCCGGGGGGGGCTTGGGGAGAGTCTGTAGCGCGCGAGGCTGGGGCGCGCAAGCG 99094
QY 636 TTTTCTGTCTATCCCGGGTACGTGGAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 695
Db 99095 TTTTCTGTCTATCCCGGGTACGTGGAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 99154
QY 696 GTTCCAGGGGGGACCTGGAACATTATCTGTGACGCTGTGCGACGCGCGCGGCGACT 755
Db 99155 GTTCCAGGGGGGACCTGGAACATTATCTGTGACGCTGTGCGACGCGCGGCGGACT 99214
QY 756 CTACCGGCATCCAGCATCTCAACCCCATCAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 815
Db 99215 CTACCGGCATCCAGCATCTCAACCCCATCAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 99274
QY 816 TAGAGATCGTGACTCCGGGGCCCAAGGTACCGGCAATGCGGGCCCTGACGCTGCGCAACTT 875
```

```
Db 99275 TAGAGATCGTGACTCCGGGGCCCAAGGTACCGGCAATGCGGGCCCTGACGCTGCGCAACTT 99334
QY 876 CTGTGCTGGCAGAGAAGCTGAACAAAGTGAAGTGAACAAGCAGCAGCAGCAGTACTGGACAC 935
Db 99335 CTGTGCTGGCAGAGAAGCTGAACAAAGTGAAGTGAACAAGCAGCAGCAGCAGTACTGGACAC 99394
QY 936 TGGCATCTCTTCCACAGGAGG 958
Db 99395 TGGCATCTCTTCCACAGGAGG 99417

RESULT 7
LOCUS AP003459 28000 bp DNA linear PRI 08-MAR-2002
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-211H5,
complete sequence.
ACCESSION AP003459
VERSION AP003459.2 GI:19263045
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-211H5.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 28000)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Mar 7, 2002 this sequence version replaced gi:13488920.
FEATURES
source
1..28000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-211H5"
BASE COUNT 5938 a 6083 c 6679 g 8300 t
ORIGIN
Query Match 25.4%; Score 724; DB 9; Length 28000;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2079 GCATGCTACAATTCGTGTGGCCATCCCGCAGCGCGCTCAAGCATCGACATCGCCA 2138
Db 3473 GCATGCTACAATTCGTGTGGCCATCCCGCAGCGCGCTCAAGCATCGACATCGCCA 3532
QY 2139 GCGCGTTTACAAGGGCTCATCGGGATGACAACCTACCTGGCTCTGAAGAACAGGCAAG 2198
Db 3533 GCGCGTTTACAAGGGCTCATCGGGATGACAACCTACCTGGCTCTGAAGAACAGGCAAG 3592
QY 2199 CAAGTACCTGCTCAACGGGCATTTCGTGTGTCGGGTGGAGCGGAGCCTGGTGTGAA 2258
Db 3593 CAAGTACCTGCTCAACGGGCATTTCGTGTGTCGGGTGGAGCGGAGCCTGGTGTGAA 3652
QY 2259 GGGCAGTCTGCTCGGTACAGCGGCACAGCGGTGGAGAGCCTGACAGCTTCCCG 2318
Db 3653 GGGCAGTCTGCTCGGTACAGCGGCACAGCGGTGGAGAGCCTGACAGCTTCCCG 3712
QY 2319 GCGCATCTGGAGCGCTGACCGTGAAGTCTCTCCGTGGGGAAAGATGACACGCCGCCG 2378
Db 3713 GCGCATCTGGAGCGCTGACCGTGAAGTCTCTCCGTGGGGAAAGATGACACGCCGCCG 3772
```

```
Oy 2379 GGTCCGCTACTCTCTATCTGCCAAAGAGCTCGGAGAGCAAGTCTCTCATCCCAA 2438
Db 3773 GGTCCGCTACTCTCTATCTGCCAAAGAGCTCGGAGAGCAAGTCTCTCATCCCAA 3832
Oy 2439 GGACCCCGGGGACCCCTCTCTGTGCACACACGGTCCCTCAGCCTCTCCAAACAGGTGA 2498
Db 3833 GGACCCCGGGGACCCCTCTCTGTGCACACACGGTCCCTCAGCCTCTCCAAACAGGTGA 3892
Oy 2499 GCAGCCGGAGCAGACGGCCCTCTCAGCCTGGGTGGCTGGCAGCTGGGGGCGCTGTCTCCGC 2558
Db 3893 GCAGCCGGAGCAGACGGCCCTCTCAGCCTGGGTGGCTGGCAGCTGGGGGCGCTGTCTCCGC 3952
Oy 2559 GAGCTCGGCGAGTGGCTCGCAGAAAGCGGGCGGTGGAGCTGTGGGGCTCGCCGGGAGCGG 2618
Db 3953 GAGCTCGGCGAGTGGCTCGCAGAAAGCGGGCGGTGGAGCTGTGGGGCTCGCCGGGAGCGG 4012
Oy 2619 CACGGTCCCTGCTGTGATCAGCCATCGCCGCTGGAGACACAAAGCTTGGGGGAGCGG 2678
Db 4013 CACGGTCCCTGCTGTGATCAGCCATCGCCGCTGGAGACACAAAGCTTGGGGGAGCGG 4072
Oy 2679 CTGCCCCACCTGGAGCTCAGCGCCTGGTCAACCTGCTCCAGAGCTGGCGCGGGGGAATT 2738
Db 4073 CTGCCCCACCTGGAGCTCAGCGCCTGGTCAACCTGCTCCAGAGCTGGCGCGGGGGAATT 4132
Oy 2739 TCAGAGCGCTCACTCAAGTGTGTGGCCACGAGCGCCGCTCTGCGCCGGGACAGTG 2798
Db 4133 TCAGAGCGCTCACTCAAGTGTGTGGCCACGAGCGCCGCTCTGCGCCGGGACAGTG 4192
Oy 2799 CAACCTGCACCGCAAGCCCGCAGGAGCTGGACTTCTGCTGCTCAGGCGCGTGTCTGA 2853
Db 4193 CAACCTGCACCGCAAGCCCGCAGGAGCTGGACTTCTGCTGCTCAGGCGCGTGTCTGA 4247

RESULT 8
AC025130
LOCUS Homo sapiens chromosome 11 clone RP11-211H6 map 11, WORKING DRAFT
DEFINITION AC025130
AC025130
VERSION AC025130.2 GI:7387384
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 172905)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
JOURNAL Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
O'Neil,D., Oliver,T.M., Norman,C.H., O'Connor,T., Peterson,K., Pierre,N.,
Riley,R., Rogov,P., Rothman,D.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Unpublished
2 (bases 1 to 172905)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
JOURNAL Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
O'Neil,D., Oliver,T.M., Norman,C.H., O'Connor,T., Peterson,K., Pierre,N.,
Riley,R., Rogov,P., Rothman,D.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Unpublished
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
```

```
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 172905)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
JOURNAL Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 3, 2000 this sequence version replaced gi:7158941.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7808
Center clone name: 211_H_6
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161899 bases at least Q40
Consensus quality: 167402 bases at least Q30
Consensus quality: 169801 bases at least Q20
Insert size: 177000; agarose-ff
Quality coverage: 4.1 in Q20 bases; agarose-ff
Quality coverage: 4.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 21: contig of 21 bp in length
2 121: gap of 100 bp
122 1343: contig of 1222 bp in length
1344 1443: gap of 100 bp
1444 2636: contig of 1193 bp in length
2637 2736: gap of 100 bp
2737 4052: contig of 1316 bp in length
4053 4152: gap of 100 bp
4153 8044: contig of 3892 bp in length
8045 8144: gap of 100 bp
8145 14949: contig of 6805 bp in length
14950 15049: gap of 100 bp
15050 23535: contig of 8486 bp in length
23536 23635: gap of 100 bp
23636 30261: contig of 6626 bp in length
```

* arbitrary, gaps between the conigs are represented as:

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1856: contig of 1856 bp in length
 * 1957 1956: gap of unknown length
 * 1957 7065: contig of 5109 bp in length
 * 7066 7165: gap of unknown length
 * 7166 21089: contig of 13924 bp in length
 * 21090 21189: gap of unknown length
 * 21190 34145: contig of 12956 bp in length
 * 34146 34245: gap of unknown length
 * 34246 48562: contig of 14317 bp in length
 * 48563 48662: gap of unknown length
 * 48663 109279: contig of 60517 bp in length
 * 109180 109279: gap of unknown length
 * 109280 170682: contig of 61403 bp in length.

FEATURES

Location/Qualifiers
 1. 170682
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /clone="RP11-121M22"
 /clone_lib="RPCI human BAC library 11"
 1. 1856
 /note="assembly_name:Contig25"
 1957. 7065
 /note="assembly_name:Contig26"
 7166. 21089
 /note="assembly_name:Contig27"
 21190. 34145
 /note="assembly_name:Contig28"
 34246. 48562
 /note="assembly_name:Contig29
 clone_end:SP6"
 48663. 109179
 /note="assembly_name:Contig30"
 109280. 170682
 /note="assembly_name:Contig31
 clone_end:T7"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

BASE COUNT 43947 a 39732 c 39514 g 46884 t 605 others

ORIGIN

Query Match 10.1% Score 288; DB 2: Length 170682;
 Best Local Similarity 100.0%; Pred. No. 3.9e-146;
 Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1258 GGTGACTGCTCTGGACCAACCAAGCCATCTCCCTGCGGAGGATCTGCCGGG 1317
 |||||
 DB 44571 GGTGACTGCTCTGGACCAACCAAGCCATCTCCCTGCGGAGGATCTGCCGGG 44512
 |||||
 QY 1318 GCCAGCTACACCTGAGCCAGCAGTGGAGTGGCTTTTGGCGTGGGCTCAAGCCCTGT 1377
 |||||
 DB 44511 GCCAGCTACACCTGAGCCAGCAGTGGAGTGGCTTTTGGCGTGGGCTCAAGCCCTGT 44452
 |||||
 QY 1378 CCTTACATGCACTGTCACCAAGCTGTGTGTCACCGGAGGCAAGGACAGATGGT 1437
 |||||
 DB 44451 CCTTACATGCACTGTCACCAAGCTGTGTGTCACCGGAGGCAAGGACAGATGGT 44392
 |||||
 QY 1438 TGCCAGACCCGCACTTCCCTGGGCGGATGGCCACCAAGCTGTGGCGAGGCAAGCTCTGC 1497
 |||||
 DB 44391 TGCCAGACCCGCACTTCCCTGGGCGGATGGCCACCAAGCTGTGGCGAGGCAAGCTCTGC 44332
 |||||
 QY 1498 CTCRAAGGGGCTGGTGGAGAGACACACCTCAACAGACAGGCTG 1545
 |||||
 DB 44331 CTCRAAGGGGCTGGTGGAGAGACACACCTCAACAGACAGGCTG 44284
 |||||

RESULT 10

AC101990 AC101990 182656 bp DNA linear HTG 21-AUG-2002
 LOCUS Mus musculus clone RP24-371J2, WORKING DRAFT SEQUENCE, 16 unordered
 DEFINITION pieces.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC101990
 AC101990.2 GI:22381363
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 house mouse.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1 (bases 1 to 182656)
 Birren,B., Nussbaum,C. and Lander,E.
 Mus musculus, clone RP24-371J2
 Unpublished

2 (bases 1 to 182656)
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., Lakocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,W.
 Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 182656)
 Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Miura,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,W.
 Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17060766.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L17833
 Center clone name: 371_J-2
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 178030 bases at least Q40
 Consensus quality: 179964 bases at least Q30
 Consensus quality: 180647 bases at least Q20
 Insert size: 172000; agarose-1p
 Insert size: 181156; sum-of-contigs
 Quality coverage: 8.1 in Q20 bases; agarose-1p
 Quality coverage: 7.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      361: contig of 361 bp in length
*      362 461: gap of 100 bp
*      462 1538: contig of 1077 bp in length
*      1539 1638: gap of 100 bp
*      1639 3397: contig of 1759 bp in length
*      3398 3497: gap of 100 bp
*      3498 4947: contig of 1450 bp in length
*      4948 5047: gap of 100 bp
*      5048 6893: contig of 1846 bp in length
*      6894 6993: gap of 100 bp
*      6994 10395: contig of 3402 bp in length
*      10396 10495: gap of 100 bp
*      10496 12584: contig of 2089 bp in length
*      12585 12684: gap of 100 bp
*      12685 14977: contig of 2293 bp in length
*      14978 15077: gap of 100 bp
*      15078 18181: contig of 3104 bp in length
*      18182 18281: gap of 100 bp
*      18282 22654: contig of 4373 bp in length
*      22655 22754: gap of 100 bp
*      22755 31261: contig of 8507 bp in length
*      31262 31361: gap of 100 bp
*      31362 43587: contig of 12226 bp in length
*      43588 43687: gap of 100 bp
*      43688 57612: contig of 13925 bp in length
*      57613 57712: gap of 100 bp
*      57713 77467: contig of 19755 bp in length
*      77468 77567: gap of 100 bp
*      77568 108724: contig of 31157 bp in length
*      108725 108824: gap of 100 bp
*      108825 182656: contig of 73832 bp in length.

```

FEATURES

source

```

Location/Qualifiers
1. .182656
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-371J2"
/clone.lib="RP24-371J2"
Male Mouse BAC"
misc_feature 1. .361
/notes="assembly_fragment"
misc_feature 462. .1538
/notes="assembly_fragment"
misc_feature 1639. .3397
/notes="assembly_fragment"
misc_feature 3498. .4947
/notes="assembly_fragment"
misc_feature 5048. .6893
/notes="assembly_fragment"
misc_feature 6994. .10395
/notes="assembly_fragment"
misc_feature 10496. .12584
/notes="assembly_fragment"
misc_feature 12685. .14977
/notes="assembly_fragment"
misc_feature 15078. .18181
/notes="assembly_fragment"

```

```

misc_feature 18282. .22654
/notes="assembly_fragment"
misc_feature 22755. .31261
/notes="assembly_fragment"
misc_feature 31362. .43587
/notes="assembly_fragment"
misc_feature 43688. .57612
/notes="assembly_fragment"
misc_feature 57713. .77467
/notes="assembly_fragment"
misc_feature 77568. .108724
/notes="assembly_fragment"
misc_feature 108825. .182656
/notes="assembly_fragment"

BASE COUNT 51468 a 41160 c 39770 g 48758 t 1500 others
ORIGIN

Query Match      2.6%  Score 74;  DB 2;  Length 182656;
Best Local Similarity 100.0%;  Pred. No. 1.4e-28;
Matches 74;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1009 ACCATGTGTGACCCCAAGAGAGAGTCTCTGTCATTGAGGACGATGGCTTCCATCAGCC 1068
|||||
Db 16446 ACCATGTGTGACCCCAAGAGAGTCTCTGTCATTGAGGACGATGGCTTCCATCAGCC 16505
|||||

QY 1069 TTCACCACTGCCCA 1082
|||||
Db 16506 TTCACCACTGCCCA 16519
|||||

RESULT 11
AC126507/c
LOCUS
DEFINITION
AC126507
VERSION
AC126507.1 GI:21700463
KEYWORDS
HTG: HTGS_PHASE1.
SOURCE
Rattus norvegicus
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 178764)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollius,B.,
Homs,F., Howard,J., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newtonson,J., Newton,N., Nguyen,N., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

```

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojubokan, I., Roife-Martin, J., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, X.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G. and Gibbs, R.
 Direct Submission
 2 (bases 1 to 178764)
 Unpublished
 Worley, K.C.
 Direct Submission
 Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 178764)
 Direct Submission
 Worley, K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information -----
 Center project name: G2PU
 Center clone name: CH230-254N12
 ----- Summary Statistics -----
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 124573 bases at least Q40
 Consensus quality: 131850 bases at least Q30
 Consensus quality: 137700 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 49 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 *
 * 1037: contig of 1037 bp in length
 * 1137: gap of unknown length
 * 1138: contig of 1113 bp in length
 * 2251: gap of unknown length
 * 2351: contig of 1229 bp in length
 * 3579: gap of unknown length
 * 3680: contig of 1195 bp in length
 * 4874: gap of unknown length
 * 4975: contig of 1551 bp in length
 * 6525: gap of unknown length
 * 6626: contig of 1280 bp in length
 * 7905: gap of unknown length
 * 8005: gap of unknown length
 * 8006: contig of 1719 bp in length
 * 9725: gap of unknown length
 * 9825: contig of 1618 bp in length
 * 11442: gap of unknown length
 * 11543: contig of 1358 bp in length
 * 12901: gap of unknown length
 * 13001: contig of 1150 bp in length
 * 14151: gap of unknown length
 * 14250: contig of 1315 bp in length
 * 15565: gap of unknown length
 * 15566: contig of 1810 bp in length
 * 17475: gap of unknown length
 * 17575: contig of 1115 bp in length
 * 18690: gap of unknown length
 * 18791: contig of 1479 bp in length
 * 20269: gap of unknown length
 * 20369: gap of unknown length
 * 20370: contig of 1558 bp in length
 * 21927: gap of unknown length
 * 21928: contig of 1558 bp in length
 * 22027: gap of unknown length
 * 22028: contig of 1971 bp in length
 * 23998: gap of unknown length
 * 23999: contig of 1581 bp in length
 * 24099: gap of unknown length
 * 25679: gap of unknown length
 * 25680: contig of 1649 bp in length
 * 25780: gap of unknown length
 * 27429: contig of 2149 bp in length
 * 27529: gap of unknown length
 * 29677: contig of 2149 bp in length
 * 29678: gap of unknown length
 * 29679: contig of 1100 bp in length
 * 30877: gap of unknown length
 * 30878: contig of 1854 bp in length
 * 30977: gap of unknown length
 * 32831: contig of 2488 bp in length
 * 32832: gap of unknown length
 * 32931: contig of 2488 bp in length
 * 32932: gap of unknown length
 * 32933: contig of 1874 bp in length
 * 33519: gap of unknown length
 * 33520: contig of 1874 bp in length
 * 37393: gap of unknown length
 * 37394: contig of 2342 bp in length
 * 37494: gap of unknown length
 * 39835: gap of unknown length
 * 39836: contig of 2854 bp in length
 * 39835: gap of unknown length
 * 42789: gap of unknown length
 * 42790: contig of 3757 bp in length
 * 42889: gap of unknown length
 * 46646: gap of unknown length
 * 46647: contig of 2442 bp in length
 * 46747: gap of unknown length
 * 49189: contig of 4517 bp in length
 * 49288: gap of unknown length
 * 53805: gap of unknown length
 * 53806: contig of 2297 bp in length
 * 53906: gap of unknown length
 * 56202: contig of 4088 bp in length
 * 56302: gap of unknown length
 * 56303: contig of 3038 bp in length
 * 58719: gap of unknown length
 * 58819: contig of 4085 bp in length
 * 62904: gap of unknown length
 * 62905: contig of 2928 bp in length
 * 65932: gap of unknown length
 * 65933: contig of 4088 bp in length
 * 66032: gap of unknown length
 * 66033: contig of 3038 bp in length
 * 70121: gap of unknown length
 * 70221: contig of 4108 bp in length
 * 73358: gap of unknown length
 * 73359: contig of 2579 bp in length
 * 77466: gap of unknown length
 * 77567: contig of 5342 bp in length
 * 80546: gap of unknown length
 * 80646: contig of 4821 bp in length
 * 85987: gap of unknown length
 * 85988: contig of 6910 bp in length
 * 86088: gap of unknown length
 * 90909: contig of 5484 bp in length
 * 91009: gap of unknown length
 * 97918: contig of 5484 bp in length
 * 98018: gap of unknown length
 * 103502: contig of 8968 bp in length
 * 103503: gap of unknown length
 * 103603: contig of 8968 bp in length
 * 112570: gap of unknown length
 * 112571: contig of 5254 bp in length
 * 112671: gap of unknown length
 * 117925: contig of 6995 bp in length
 * 117925: gap of unknown length
 * 118025: contig of 8968 bp in length
 * 125020: gap of unknown length
 * 125119: contig of 9664 bp in length
 * 125120: gap of unknown length
 * 134087: contig of 7748 bp in length
 * 134187: gap of unknown length
 * 134188: contig of 7027 bp in length
 * 143852: gap of unknown length
 * 143852: contig of 7027 bp in length
 * 143952: gap of unknown length
 * 151699: contig of 9596 bp in length
 * 151700: gap of unknown length
 * 151700: contig of 9596 bp in length
 * 158000: gap of unknown length
 * 158000: contig of 9596 bp in length
 * 158827: gap of unknown length
 * 158827: contig of 9596 bp in length
 * 158927: gap of unknown length
 * 158927: contig of 9596 bp in length

```

* 168523 168622: gap of unknown length
* 168623 178764: contig of 10142 bp in length.
FEATURES
  Location/Qualifiers
    1..178764
      /organism="Rattus norvegicus"
      /db_xref="taxon:10116"
      /clone="CH230-254N12"
BASE COUNT 47487 a 39119 c 38897 g 45423 t 7838 others
ORIGIN
  Query Match
  Best Local Similarity 100.08; Pred. No. 2e-21;
  Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1942 GTCTGTGTCACGAAGTCATCAAGCTGGCTGTGTGATGGGAACCTGGGCTCCAAGAAG 2001
      |||||
Db 150003 GTCTGTGTCACGAAGTCATCAAGCTGGCTGTGTGATGGGAACCTGGGCTCCAAGAAG 149944
QY 2002 A 2002
Db 149943 A 149943

RESULT 12
AC126507
LOCUS      178764 bp      DNA      linear      HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-254N12, *** SEQUENCING IN PROGRESS
AC126507
VERSION    AC126507.1 GI:21700463
KEYWORDS   HTG; HTGS-PHASE1.
SOURCE     Rattus norvegicus.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE
  1 (bases 1 to 178764)
    Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
    Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
    Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
    Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
    Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
    Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
    Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
    Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.F., David, R.,
    Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
    Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
    Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
    Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
    Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
    Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
    Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
    Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
    Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
    Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
    Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
    Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
    Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
    Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
    Lozador, R.J., Lu, X., Lucier, A., Lucier, R., Luna, P., Ma, J.,
    Maheshwari, M., Mapua, P., Martindale, A., Martinez, E.,
    Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
    Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
    Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
    Nguyen, N., Nickerson, E., Nwokenwo, S., Ogwu, M., Okwuonu, G.,
    Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
    Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
    Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
    Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
    Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
    Sutton, A., Svatek, A., Tabar, P., Tamerisa, A., Tamerisa, K., Tang, H.,
    Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
    Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,

```

Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 178764)
Worley, K.C.
Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Worley, K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZFU
Center clone name: CH230-254N12
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124573 bases at least Q40
Consensus quality: 131850 bases at least Q30
Consensus quality: 137700 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.


```
* 21928 22027: gap of unknown length
* 22028 23998: contig of 1971 bp in length
* 23999 24098: gap of unknown length
* 24099 25679: contig of 1581 bp in length
* 25680 25779: gap of unknown length
* 25780 27428: contig of 1649 bp in length
* 27429 27528: gap of unknown length
* 27529 29677: contig of 2149 bp in length
* 29678 29777: gap of unknown length
* 29778 30877: contig of 1100 bp in length
* 30878 30977: gap of unknown length
* 30978 32831: contig of 1854 bp in length
* 32832 32931: gap of unknown length
* 32932 35419: contig of 2488 bp in length
* 35420 35519: gap of unknown length
* 35520 37393: contig of 1874 bp in length
* 37394 37493: gap of unknown length
* 37494 39835: contig of 2342 bp in length
* 39836 39935: gap of unknown length
* 39936 42789: contig of 2854 bp in length
* 42790 42889: gap of unknown length
* 42890 46646: contig of 3757 bp in length
* 46647 46746: gap of unknown length
* 46747 49188: contig of 2442 bp in length
* 49189 49288: gap of unknown length
* 49289 53805: contig of 4517 bp in length
* 53806 53905: gap of unknown length
* 53906 56202: contig of 2297 bp in length
* 56203 56302: gap of unknown length
* 56303 58719: contig of 2417 bp in length
* 58720 58819: gap of unknown length
* 58820 62904: contig of 4085 bp in length
* 62905 63004: gap of unknown length
* 63006 65932: contig of 2928 bp in length
* 65933 66032: gap of unknown length
* 66033 70120: contig of 4088 bp in length
* 70121 70220: gap of unknown length
* 70221 73258: contig of 3038 bp in length
* 73259 73358: gap of unknown length
* 73359 77466: contig of 4108 bp in length
* 77467 77566: gap of unknown length
* 77567 80545: contig of 2979 bp in length
* 80546 80646: gap of unknown length
* 80646 85987: contig of 5342 bp in length
* 85988 86087: gap of unknown length
* 86088 90908: contig of 4821 bp in length
* 90909 91008: gap of unknown length
* 91009 97918: contig of 6910 bp in length
* 97919 98018: gap of unknown length
* 98019 103502: contig of 5484 bp in length
* 103503 103602: gap of unknown length
* 103603 112570: contig of 8968 bp in length
* 112571 112670: gap of unknown length
* 112671 117924: contig of 5254 bp in length
* 117925 118024: gap of unknown length
* 118025 125013: contig of 6995 bp in length
* 125014 125119: gap of unknown length
* 125120 134087: contig of 8968 bp in length
* 134088 134187: gap of unknown length
* 134188 143851: contig of 9664 bp in length
* 143852 143951: gap of unknown length
* 143952 151699: contig of 7748 bp in length
* 151700 151799: gap of unknown length
* 151800 158826: contig of 7027 bp in length
* 158827 158926: gap of unknown length
* 158927 168522: contig of 9596 bp in length
* 168523 168622: gap of unknown length
* 168623 178764: contig of 10142 bp in length.
*
* Location/Qualifiers
* 1..178764
* /organism="Rattus norvegicus"
* /db_xref="taxon:10116"
* /clone="CH230-254N12"
*
FEATURES
source
    QUERY MATCH
    Best Local Similarity 100.0%; Score 44; DB 2; Length 178764;
    Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 511 ACCTCTCGCTCGGGGTGCGCTCGGCTGGAACCCCGCATCCT 554
    Db 101958 ACCTCTCGCTCGGGGTGCGCTCGGCTGGAACCCCGCATCCT 102001
    RESULT 13
    AC101990/c
    LOCUS
    DEFINITION
        Mus musculus clone RP24-371J2, WORKING DRAFT SEQUENCE, 16 unordered
        pieces.
    ACCESSION
        AC101990
    VERSION
        AC101990.2 GI:22381363
    KEYWORDS
        HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
    SOURCE
        house mouse.
    ORGANISM
        Mus musculus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    REFERENCE
        1 (bases 1 to 182656)
        Birren,B., Nusbaum,C. and Lander,E.
    TITLE
        Mus musculus, clone RP24-371J2
    JOURNAL
        Unpublished
    REFERENCE
        2 (bases 1 to 182656)
        Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
        Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
        Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
        Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
        Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
        Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
        Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
        Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
        Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
        Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
        MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
        McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
        Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
        Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
        Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
        Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
        Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
        Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
        Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
        Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
        Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
        Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
    Direct Submission
    Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
    3 (bases 1 to 182656)
    TITLE
        JOURNAL
    REFERENCE
        AUTHORS
        Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
        Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
        Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
        Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
        Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
        Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
        Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
        Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
        Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
        McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
        Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
        O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
        Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
        Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
        Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
        Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
        Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
        Zembek,L., Zimmer,A. and Zody,M.
```


Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK plate: 8 Row: j Column: 3

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

Location/Qualifiers

1..3927

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="CZECH II"

/clone="IMAGE:3491991"

/tissue_type="Mammary tumor metastatized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

<1..1024

/codon_start=2

/product="Similar to a disintegrin and metalloproteinase with thrombospondin motifs 1 (ADAMTS-1)"

/protein_id="AAH09667.1"

/db_xref="GI:16307156"

/translation="PRVRKCLICRANGTYFYVLAPKVVDTGLCTPDSTVCVQKCIKACDGNLGSKKFDKCGVCGDNKSKRVYGLTKPMHGYNFVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVGSVLRYSGTGTAVESLQASRPILFLEPTVLSVGMTPPRVSYFLPEKPEDKSTRPKDPRGSLVLRNSVLSLSNOVEOPDNRPAPRWAGSWGPCSVCSGLQKRAVDGRDSDPGQGSACDVRHPLFKRACGPEPTWELGNWSPCSKSGRGFKRRPLKCVGHGRLRLARDQCDLRRKPQELDFCVLRPC"

BASE COUNT 938 a 920 c 1093 g 976 t

ORIGIN

Query Match 1.2%; Score 35; DB 10; Length 3927;

Best Local Similarity 100.0%; Pred. No. 5.3e-07;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1942 GTCTGTCTCCAAAGCAAGTCATCAAGGCTGGCTG 1976

|||||

Db 113 GTCTGTCTCCAAAGCAAGTCATCAAGGCTGGCTG 147

Search completed: May 16, 2003, 03:53:15

Job time : 10280 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 22:55:17 ; Search time 577 Seconds
(without alignments)
11135.097 Million cell updates/sec

Title: US-09-965-631-3
Perfect score: 2853
Sequence: 1 atgctctgtggcatcct.....gcgtctgaggcgtgctga 2853

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 24

Total number of hits satisfying chosen parameters: 19

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2853	100.0	2853	24	Human protease cDN
2	2853	100.0	3446	24	Human protease cDN
3	2695	94.5	2853	22	Human metalloprote
4	2196	77.0	2930	24	Human metalloprote
5	2196	77.0	2937	24	Human metalloprote
6	1091	38.2	1104	24	Human protease cDN
7	959	33.6	966	24	Human protease cDN
8	481	16.9	1143	21	Human metalloprote
9	50	1.8	1518	21	Rat metalloprotein

c	10	30	1.1	30	22	AAH41029	Adaptor primer SEQ
c	11	30	1.1	30	22	AAH41030	Adaptor primer SEQ
c	12	30	1.1	41	22	AAH41017	PCR primer specifi
c	13	27	0.9	27	22	AAH41018	PCR primer specifi
c	14	27	0.9	37	22	AAH41019	PCR primer specifi
c	15	27	0.9	38	22	AAH41024	PCR primer specifi
c	16	27	0.9	38	22	AAH41033	PCR primer for met
c	17	27	0.9	38	22	AAH41034	PCR primer for met
c	18	25	0.9	610	24	ABQ44966	Oligonucleotide fo
c	19	25	0.9	610	24	ABQ44967	Oligonucleotide fo

ALIGNMENTS

RESULT 1

AAD35569

ID AAD35569 standard; cDNA; 2853 BP.

AC AAD35569;

XX 26-JUL-2002 (first entry)

DE Human protease cDNA #2.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;
KW high blood pressure; arthritis; connective tissue disorder; infertility;
KW gene therapy; enzyme; gene; ss.
XX Homo sapiens.

XX Key

XX Location/Qualifiers
FT CDS
FT 1..2853
FT /*tag= a
FT /product= "Human protease #2"

WO200226949-A2.

PD 04-APR-2002.

PF 27-SEP-2001; 2001WO-US30350.

PR 29-SEP-2000; 2000US-236689P.

PA (LEXI-) LEXICON GENETICS INC.

PI Friddle CJ, Hilbun E;

DR WPI: 2002-372123/40.

DR P-PSDB; AAD22541.

XX Novel nucleic acid encoding a human protease, useful as a hybridization
PT probe for screening libraries and assessing gene expression patterns -
XX Claim 1; Page 35-36; 41pp; English.

XX The present sequence is a cDNA encoding novel human protein (NHP),
CC human protease. NHPs share structural similarity with animal proteases
CC particularly zinc metalloproteases. Sequences of the invention are
CC useful in therapeutic, diagnostic and pharmacogenomic applications.
CC NHP polynucleotides are used as hybridisation probes for screening
CC libraries and assessing gene expression patterns. They can also be
CC used for treating related biological disorders such as obesity, high
CC blood pressure, arthritis, connective tissue disorders and infertility.
CC They are also used in gene therapy.

XX Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;

Query Match 100.0%; Score 2853; DB 24; Length 2853;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGCATCTCAACCTGGCTTTCGCGGGGGAACGCTGAGGCTCTGAG 60

Db	1		ATGCTTCTGCTGGGCACTCTTAACCTTGCCTTTCGCGGGCAGCCGCTGAGGCTCTGAG	60
Qy	61		CCAGAGCGGAGGTAGTCTGTTCCCATCCGACTGACCCCGGACATTAACAGCGCGCGCTAC	120
Db	61		CCAGAGCGGAGGTAGTCTGTTCCCATCCGACTGACCCCGGACATTAACAGCGCGCGCTAC	120
Qy	121		TACTGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCACAGCATTT	180
Db	121		TACTGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCACAGCATTT	180
Qy	181		CAGGAGGACTTTTACCTACACTGACCGCGATGCTCAGTCTTGGCTCCGCGCTCTCC	240
Db	181		CAGGAGGACTTTTACCTACACTGACCGCGATGCTCAGTCTTGGCTCCGCGCTCTCC	240
Qy	241		ACTGAGCATCTGGCGTCCCTCCAGGGCTCAACGGGGCTCTCAGACCTTGCAGCGC	300
Db	241		ACTGAGCATCTGGCGTCCCTCCAGGGCTCAACGGGGCTCTCAGACCTTGCAGCGC	300
Qy	301		TGCTTCTATTCTGGGACGTGAACCCGAGCGGACTGTTCCGCTGTGAGCCTGTGC	360
Db	301		TGCTTCTATTCTGGGACGTGAACCCGAGCGGACTGTTCCGCTGTGAGCCTGTGC	360
Qy	361		GGGGGCTCCGCGGACCTTTGGCTACCGAGCGCGGAGTGTGATTAGCCCGCTGCC	420
Db	361		GGGGGCTCCGCGGAGCCTTTGGCTACCGAGCGCGGAGTGTGATTAGCCCGCTGCC	420
Qy	421		AATGCTAGCGCGCGCGGCGAGCGCAACGCCAGGGCGCACACTTCTCCAGCGCGG	480
Db	421		AATGCTAGCGCGCGCGGCGAGCGCAACGCCAGGGCGCACACTTCTCCAGCGCGG	480
Qy	481		GGTGTTCGGGGCGGCTTCGGGAGACCCACCTCTCGCTCGGCTCGGGCTCGGGCTG	540
Db	481		GGTGTTCGGGGCGGCTTCGGGAGACCCACCTCTCGCTCGGCGTGGGCTCGGGCTG	540
Qy	541		AACCCGCCATCTACGGGCGCTTGAACCTTAAAGCCGCGCGGGGCTTCGGGAG	600
Db	541		AACCCGCCATCTACGGGCGCTTGAACCTTAAAGCCGCGCGGGGCTTCGGGAG	600
Qy	601		AGTCGTAGCGGCGGAGTCTGGGCGGCCAGCGTTCGTGTCTATCCCGGCTAGTG	660
Db	601		AGTCGTAGCGGCGGAGTCTGGGCGGCCAGCGTTCGTGTCTATCCCGGCTAGTG	660
Qy	661		GAGAGCTGTGTGCGGGACGAGTCAATGTTCAAGTTCACGGCGGGAGCTGGACAT	720
Db	661		GAGAGCTGTGTGCGGGACGAGTCAATGTTCAAGTTCACGGCGGGAGCTGGACAT	720
Qy	721		TATCTGTGAGCTGTGTGCGAAGCGGCGGCGACTCTACCGCCATCCAGCATCCTCAAC	780
Db	721		TATCTGTGAGCTGTGTGCGAAGCGGCGGCGACTCTACCGCCATCCAGCATCCTCAAC	780
Qy	781		CCCATCAACATCGTTGTGTCAGAGTGTCTTCTAGAGATCTGACTCGGGCGCCAAAG	840
Db	781		CCCATCAACATCGTTGTGTCAGAGTGTCTTCTAGAGATCTGACTCGGGCGCCAAAG	840
Qy	841		GTCCAGCGCAATGGCGCCTTGACGCTGCGCAACTTCTGTGCTGCGAGAAAGCTGAAC	900
Db	841		GTCCAGCGCAATGGCGCCTTGACGCTGCGCAACTTCTGTGCTGCGAGAAAGCTGAAC	900
Qy	901		AAAGTGTGACAGACACCCGAGTACTGGGACACTGCCATCTCTTACCAGCAGGAC	960
Db	901		AAAGTGTGACAGACACCCGAGTACTGGGACACTGCCATCTCTTACCAGCAGGAC	960
Qy	961		CTGTGTGAGCCACCACTGTGACACCTTGGGCGATGGCTGTGCTGCTGCTGCTGAC	1020
Db	961		CTGTGTGAGCCACCACTGTGACACCTTGGGCGATGGCTGTGCTGCTGCTGCTGAC	1020
Qy	1021		CCCAAGAAAGCTGCTGTGATTTAGGACGATGGCTTCCTTACCACCTGCC	1080
Db	1021		CCCAAGAAAGCTGCTGTGATTTAGGACGATGGCTTCCTTACCACCTGCC	1080
Qy	1081		CACGAGCTGGGCCACGTGTTCAACATGCCCCATGACAATGTGAATCTGTGAGAGTG	1140

D	b	1081	C	A	G	A	G	T	G	G	G	C	A	G	T	G	T	T	A	A	C	A	T	G	C	C	C	A	T	G	A	C	A	A	T	G	T	G	A	A	A	G	T	C	T	G	A	G	A	G	T	G	1140
Q	y	1141	T	T	T	G	G	A	G	T	C	C	G	A	C	A	A	C	A	T	G	A	T	G	T	C	C	C	G	C	A	C	C	T	C	A	T	C	C	A	G	A	T	C	G	A	C	C	T	G	C	1200	
D	b	1141	T	T	T	G	G	A	G	A	G	T	C	C	G	A	C	A	A	C	A	T	G	A	T	G	T	C	C	C	G	A	C	C	T	C	A	T	C	C	A	G	A	T	C	G	A	C	C	T	G	1200	
Q	y	1201	A	A	C	C	C	T	G	A	G	T	C	C	A	T	A	C	A	T	A	C	A	T	A	C	C	G	A	C	T	T	C	T	G	A	C	A	C	G	A	C	G	A	C	G	A	C	G	T	1260		
D	b	1201	A	A	C	C	C	T	G	A	G	T	C	C	A	T	A	C	A	T	A	C	A	T	A	C	C	G	A	C	T	T	C	T	G	A	C	A	C	G	A	C	G	A	C	G	A	C	G	T	1260		
Q	y	1261	G	A	C	T	G	C	T	C	T	G	A	C	A	C	C	C	A	G	A	A	G	C	C	A	T	C	T	C	C	T	G	C	C	G	A	G	A	T	C	T	G	C	C	G	A	C	G	T	1320		
D	b	1261	G	A	C	T	G	C	T	C	T	G	A	C	A	C	C	C	A	G	A	A	G	C	C	A	T	C	T	C	C	T	G	C	C	G	A	G	A	T	C	T	G	C	C	G	A	C	G	T	1320		
Q	y	1321	A	G	C	T	A	C	A	C	C	T	G	A	C	A	G	A	G	T	G	G	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1380			
D	b	1321	A	G	C	T	A	C	A	C	C	T	G	A	C	A	G	A	G	T	G	G	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1380				
Q	y	1381	T	A	C	A	T	G	A	G	T	C	A	C	A	A	G	T	G	T	G	T	G	T	G	C	A	C	C	G	G	A	A	G	C	A	A	G	A	G	A	C	A	G	A	T	G	T	G	1440			
D	b	1381	T	A	C	A	T	G	A	G	T	C	A	C	A	A	G	T	G	T	G	T	G	T	G	C	A	C	C	G	G	A	A	G	C	A	A	G	A	G	A	C	A	G	A	T	G	T	G	1440			
Q	y	1441	C	A	G	A	C	C	C	A	C	T	T	C	C	C	T	G	G	C	C	A	G	A	C	A	C	A	G	T	G	T	G	C	G	A	G	G	A	G	A	C	A	G	T	C	T	A	C	T	1500		
D	b	1441	C	A	G	A	C	C	C	A	C	T	T	C	C	C	T	G	G	C	C	A	G	A	C	A	C	A	G	T	G	T	G	C	G	A	G	A	G	A	C	A	G	T	C	T	A	C	T	1500			
Q	y	1501	A	A	A	G	G	C	C	T	G	G	A	G	A	C	A	C	A	C	C	T	C	A	A	A	A	G	C	A	C	A	A	G	C	A	C	A	A	G	C	A	A	G	C	A	A	G	C	1560			
D	b	1501	A	A	A	G	G	C	C	T	G	G	A	G	A	C	A	C	A	C	C	T	C	A	A	A	A	G	C	A	C	A	A	G	C	A	C	A	A	G	C	A	A	G	C	A	A	G	C	1560			
Q	y	1561	A																																																		

QY 2221 TTCGTGGTCTCGCGGTGAGCGGACCTGTGTGAAG36CAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCGTGGTCTCGCGGTGAGCGGACCTGTGTGAAG36CAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCACGGGCACAGCGGTGAGAGCCTGAGAGCTTCCCGG3CCATCCTGTGAGCGCGTGACC 2340
Db 2281 GGCACGGGCACAGCGGTGAGAGCCTGAGAGCTTCCCGG3CCATCCTGTGAGCGCGTGACC 2340
QY 2341 GTGAGGTCCTCTCCGTGGGGAAGATGACACCGCCCGGTCGCGTACTCTTCTATCTG 2400
Db 2341 GTGAGGTCCTCTCCGTGGGGAAGATGACACCGCCCGGTCGCGTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGAGGACAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
Db 2401 CCCAAGAGCCTCGGAGGACAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
QY 2461 TTGCACACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACAGACGCCCT 2520
Db 2461 TTGCACACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACAGACGCCCT 2520
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCCGTCGTCGCGGAGCTGGCGAGTGGCGCTGCAG 2580
Db 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCCGTCGTCGCGGAGCTGGCGAGTGGCGCTGCAG 2580
QY 2581 AAGCGGCGGTGGACTGTCTGGGGCTCCGCGGCGAGCGGTCCTGCTGTGATGCA 2640
Db 2581 AAGCGGCGGTGGACTGTCTGGGGCTCCGCGGCGAGCGGTCCTGCTGTGATGCA 2640
QY 2641 GCCATCGGCGGTGGAGACACAGCTGCGGGAGCGCTGCCACCTGGAGCTCAGC 2700
Db 2641 GCCATCGGCGGTGGAGACACAGCTGCGGGAGCGCTGCCACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCACCTGTCTCAAGAGCTGCGCGGCGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGGTCACCTGTCTCAAGAGCTGCGCGGCGGATTTTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCCACGAGCGCGGTGCTGCGCGGAGCACAGTGCACCTTGACCGCCAGCCCGCAG 2820
Db 2761 GTGGGCCACGAGCGCGGTGCTGCGCGGAGCACAGTGCACCTTGACCGCCAGCCCGCAG 2820
QY 2821 GAGCTGACTTCTCGCTCCCTGAGSCCGTGCTGA 2853
Db 2821 GAGCTGACTTCTCGCTCCCTGAGSCCGTGCTGA 2853

RESULT 2

AAD35571

ID AAD35571 standard; cdna; 3446 BP.

AC AAD35571;

XX AAD35571;

DT 26-JUL-2002 (first entry)

XX Human protease cdna #4.

DE Human; novel human protein; NHP; protease; biological disorder; obesity;

KW high blood pressure; arthritis; connective tissue disorder; infertility;

KW gene therapy; enzyme; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO200226949-A2.

PN 04-APR-2002.

XX 27-SEP-2001; 2001WO-US30350.

XX 29-SEP-2000; 2000US-236689P.

PR (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Hilbun E;

XX

PI

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

WPI; 2002-372123/40.

Novel nucleic acid encoding a human protease, useful as a hybridization probe for screening libraries and assessing gene expression patterns -

Disclosure; Page 40-41; 41pp; English.

The present sequence is a cDNA encoding novel human protein (NHP), human protease. NHPs share structural similarity with animal proteases particularly zinc metalloproteases. Sequences of the invention are useful in therapeutic, diagnostic and pharmacogenomic applications. NHP polynucleotides are used as hybridisation probes for screening libraries and assessing gene expression patterns. They can also be used for treating related biological disorders such as obesity, high blood pressure, arthritis, connective tissue disorders and infertility. They are also used in gene therapy.

Sequence 3446 BP; 612 A; 1114 C; 1101 G; 619 T; 0 other;

Query Match 100.0%; Score 2853; DB 24; Length 3446;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCCTTAACCCCTGGCTTTCGCGGCGGACACCCCTGGAGGCTCTGAG 60

Db 397 ATGCTTCTGCTGGGATCCTTAACCCCTGGCTTTCGCGGCGGACACCCCTGGAGGCTCTGAG 456

QY 61 CCAGAGCGGAGTAGTCTGCTTCCATCCGACTGGACCCGAGCATTAACGCGCGCGCTAC 120

Db 457 CCAGAGCGGAGTAGTCTGCTTCCATCCGACTGGACCCGAGCATTAACGCGCGCGCTAC 516

QY 121 TACTGGGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180

Db 517 TACTGGGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 576

QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCCGCTTCC 240

Db 577 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCCGCTTCC 636

QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACTGGAGCG 300

Db 637 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACTGGAGCG 696

QY 301 TGTCTTCTTCTGGGACGTGAACCGGACGCGGACTCGTTGCTGCTGAGCCTGTGC 360

Db 697 TGTCTTCTTCTGGGACGTGAACCGGACGCGGACTCGTTGCTGCTGAGCCTGTGC 756

QY 361 GGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGGAGTATGTCATTAGCCGCTGCC 420

Db 757 GGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGGAGTATGTCATTAGCCGCTGCC 816

QY 421 AATGCTAGGCGCGCGGCGCGGAGCGCAACAGCGGCGCACACTTCTTCAGCGCGCG 480

Db 817 AATGCTAGGCGCGCGGCGCGGAGCGCAACAGCGGCGCACACTTCTTCAGCGCGCG 876

QY 481 GTGTTCGGGGCGGCTTCCGAGACCCCACTCTCGCTCGGGGTGGCTTCGGGCTGG 540

Db 877 GTGTTCGGGGCGGCTTCCGAGACCCCACTCTCGCTCGGGGTGGCTTCGGGCTGG 936

QY 541 AACCCTCCCATCTTACGGGCTTGGAGCCTTACAAAGCGCGGCGGGGCTTCGGGCGAG 600

Db 937 AACCCTCCCATCTTACGGGCTTGGAGCCTTACAAAGCGCGGCGGGGCTTCGGGCGAG 996

QY 601 AGTCGTAGCGCGGCGGAGTCTGGGCGCGCAAGCGTTTCTGTCATCCCGCGTACG 660

Db 997 AGTCGTAGCGCGGCGGAGTCTGGGCGCGCAAGCGTTTCTGTCATCCCGCGTACG 1056

QY 661 GAGACGCTGGTGTGCGGAGCAGTCAATGTTCAGTTCCAGCGCGGCGGCTTGAACAT 720

Db 1057 GAGACGCTGGTGTGCGGAGCAGTCAATGTTCAGTTCCAGCGCGGCGGCTTGAACAT 1116

QY 721 TATCTGCTGACGCTGCTGGCAACGCGGCGGAGCTTACCAGCATCTCCAGCATCTCAAC 780

Db 1117 TATCTGTCAGCGTGTGGCAACGGGGCGGCACTCTACGCCATCCCAAGCATCTCTCAAC 1176
QY 781 CCATCAACATCTGTGGTCAAGGTGCTGCTTCTTAGAGATGCTGACTCGGGCCCAAG 840
Db 1177 CCATCAACATCTGTGGTCAAGGTGCTGCTTCTTAGAGATGCTGACTCGGGCCCAAG 1236
QY 841 GTACCCGGCAATGGGCGCTGAGCGTGGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 900
Db 1237 GTACCCGGCAATGGGCGCTGAGCGTGGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 1296
QY 901 AAAGTGAAGTGAAGCAACCCAGTACTGGGACACTGGCCATCTCTTCAACAGGCAAGGAC 960
Db 1297 AAAGTGAAGTGAAGCAACCCAGTACTGGGACACTGGCCATCTCTTCAACAGGCAAGGAC 1356
QY 961 CTGTGTGGGCCACCACTGTGACACCTGGGATGGCTGATGTGGTACCATGTGTGAC 1020
Db 1357 CTGTGTGGGCCACCACTGTGACACCTGGGATGGCTGATGTGGTACCATGTGTGAC 1416
QY 1021 CCCAAGAGAAGTGTCTGTCAATTGAGGAGATGGGCTTCCATCAGCCTTCAACACTGCC 1080
Db 1417 CCCAAGAGAAGTGTCTGTCAATTGAGGAGATGGGCTTCCATCAGCCTTCAACACTGCC 1476
QY 1081 CACGAGCTGGGCCAGTGTTCACATGCCCATGACACATGTGAAGTCTGTGAGGAGTG 1140
Db 1477 CACGAGCTGGGCCAGTGTTCACATGCCCATGACACATGTGAAGTCTGTGAGGAGTG 1536
QY 1141 TTTGGGAAGTCCGAGCAACCACTGATGTCGCCGACCTCATCCAGATCGACCGTGCC 1200
Db 1537 TTTGGGAAGTCCGAGCAACCACTGATGTCGCCGACCTCATCCAGATCGACCGTGCC 1596
QY 1201 AACCCCTGGTCAAGCTGAGTGTGCCATCATCACCGACTTCTCCCGAGGATCTCCCGGGGCC 1260
Db 1597 AACCCCTGGTCAAGCTGAGTGTGCCATCATCACCGACTTCTCCCGAGGATCTCCCGGGGCC 1656
QY 1261 GACTGCTCTGGACCAACCCAGCAACCCATCTCTCCCTGCCCGAGGATCTCCCGGGGCC 1320
Db 1657 GACTGCTCTGGACCAACCCAGCAACCCATCTCTCCCTGCCCGAGGATCTCCCGGGGCC 1716
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCCCT 1380
Db 1717 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCCCT 1776
QY 1381 TACATGCAAGTACTGACCAAGCTGTGTGCACCGGGAAGGCCAAGGACAGATGTTGTGC 1440
Db 1777 TACATGCAAGTACTGACCAAGCTGTGTGCACCGGGAAGGCCAAGGACAGATGTTGTGC 1836
QY 1441 CAGACCCGCCACTTCCCTGCGGCGATGGCACAGCTGTGGCGAGGCAAGCTCTGCCCTC 1500
Db 1837 CAGACCCGCCACTTCCCTGCGGCGATGGCACAGCTGTGGCGAGGCAAGCTCTGCCCTC 1896
QY 1501 AAAGGGCTCGGTGGAGAGACAACTCAACAGCACAGGATGGATGTTCTTGGGCC 1560
Db 1897 AAAGGGCTCGGTGGAGAGACAACTCAACAGCACAGGATGGATGTTCTTGGGCC 1956
QY 1561 AAATGGATTCCTATGCCCCCTGCTCGGCACATGTGGTGGGCGTGCAGCTGGCCAGG 1620
Db 1957 AAATGGATTCCTATGCCCCCTGCTCGGCACATGTGGTGGGCGTGCAGCTGGCCAGG 2016
QY 1621 AGGAGTGCACAAACCCACCCCTGCCAACCGGGGCAAGTACTGCGAGGAGTGAAGGTTG 1680
Db 2017 AGGAGTGCACAAACCCACCCCTGCCAACCGGGGCAAGTACTGCGAGGAGTGAAGGTTG 2076
QY 1681 AAATACCGATCTGATCTGGAGCCTTCCCGAGCTCAGCTCCCGGAAAGAGCTTCCGG 1740
Db 2077 AAATACCGATCTGATCTGGAGCCTTCCCGAGCTCAGCTCCCGGAAAGAGCTTCCGG 2136
QY 1741 GAGGAGCAGTGTAGGCTTTTCAACGGGTACAAACACAGCACCAACCGGCTCACTCTGCC 1800
Db 2137 GAGGAGCAGTGTAGGCTTTTCAACGGGTACAAACACAGCACCAACCGGCTCACTCTGCC 2196
QY 1801 GTGGCATGGTGGCCCAAGTACTCGGGCTGTCTCCCGGGGACAAAGTGCAGAGCTCATCTGC 1860
|||||

Db 2197 GTGGCATGGTGGCCCAAGTACTCCGGCTGTCTCCCGGGGACAAAGTGCAGAGCTCACTGC 2256
QY 1861 CGAGCAATGCACTGGCTACTTCTATGTGTGCACCCCAAGGTGGTGGAGCGCACCTG 1920
Db 2257 CGAGCAATGCACTGGCTACTTCTATGTGTGCACCCCAAGGTGGTGGAGCGCACCTG 2316
QY 1921 TGCTCTCTGACTCCACCTCGTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 2317 TGCTCTCTGACTCCACCTCGTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGAT 2376
QY 1981 GGGACCTGGGCTCCAAAGAGATTCCACAAGTGTGGGTGTGTGGGGAGACAATAAG 2040
Db 2377 GGGACCTGGGCTCCAAAGAGATTCCACAAGTGTGGGTGTGTGGGGAGACAATAAG 2436
QY 2041 AGCTCCAAAGAAGGTGACTGGACTCTTCCACCAAGCCCATGCTCAATTTCTGGTG 2100
Db 2437 AGCTCCAAAGAAGGTGACTGGACTCTTCCACCAAGCCCATGCTCAATTTCTGGTG 2496
QY 2101 GCCATCCCGGAGGGCGCTCAAGCATCGACATCGCCAGCCGGTTACAAAGGGCTGATC 2160
Db 2497 GCCATCCCGGAGGGCGCTCAAGCATCGACATCGCCAGCCGGTTACAAAGGGCTGATC 2556
QY 2161 GGGGATGACAACCTACTGCTCTGAAGAAACAGCCAAAGGCAAGTACTCTCAAGGGCAT 2220
Db 2557 GGGGATGACAACCTACTGCTCTGAAGAAACAGCCAAAGGCAAGTACTCTCAAGGGCAT 2616
QY 2221 TTTCTGGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCAAGTACTCTCTCGGTTAGC 2280
Db 2617 TTTCTGGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCAAGTACTCTCTCGGTTAGC 2676
QY 2281 GGCAGGGCACAGCGGTGGAGAGCTGCAGAGCTTCCCGGCCCATCTCTGGAGCCGTGACC 2340
Db 2677 GGCAGGGCACAGCGGTGGAGAGCTGCAGAGCTTCCCGGCCCATCTCTGGAGCCGTGACC 2736
QY 2341 GTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2737 GTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAGAGCCTCGGAGAGCAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTC 2460
Db 2797 CCCAAGAGCCTCGGAGAGCAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTC 2856
QY 2461 TTGCACAACAGCGTCTCAGCCTCTCCAAACAGTGGAGCAGCCGACAGAGGCCCT 2520
Db 2857 TTGCACAACAGCGTCTCAGCCTCTCCAAACAGTGGAGCAGCCGACAGAGGCCCT 2916
QY 2521 GCACGCTGGGTGGTGGCAGTGGGGCGGTGCTCCGGAGCTGGCGGAGTGGCCTGAG 2580
Db 2917 GCACGCTGGGTGGTGGCAGTGGGGCGGTGCTCCGGAGCTGGCGGAGTGGCCTGAG 2976
QY 2581 AAGCGGCGGTGACTGTGGGGCTCCCGGGCAGCGCACGCTCCCTGCTGTGAAGCA 2640
Db 2977 AAGCGGCGGTGACTGTGGGGCTCCCGGGCAGCGCACGCTCCCTGCTGTGAAGCA 3036
QY 2641 GCCATCGGCCGTGGAGACACAAGCTTCGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
Db 3037 GCCATCGGCCGTGGAGACACAAGCTTCGGGGAGCCCTGCCACCTGGGAGCTCAGC 3096
QY 2701 GCTGTGTACCTGTCTCCAAAGAGCTTCGGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 3097 GCTGTGTACCTGTCTCCAAAGAGCTTCGGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 3156
QY 2761 GTGGGCAAGGAGCGGCTGTGTGCCCGGGACCAAGTGCAGCTGCAGCGCAAGCCCCAG 2820
Db 3157 GTGGGCAAGGAGCGGCTGTGTGCCCGGGACCAAGTGCAGCTGCAGCGCAAGCCCCAG 3216
QY 2821 GAGCTGAGCTTCTGGCTCTGAGCGCTGCTGA 2853
|||||
Db 3217 GAGCTGAGCTTCTGGCTCTGAGCGCTGCTGA 3249
|||||

RESULT 3
AAH41003

ID	AAH41003 standard; cDNA; 2853 BP.	
XX	AAH41003;	
XX		
DT	23-AUG-2001 (first entry)	
DE	Human metalloprotease MDTs6 cDNA.	
XX		
KW	Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;	
KW	osteopathic; antiarthritic; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..2853
FT		/*tag= a
FT		/product= "Metalloprotease"
FT		/note= "Aggrecanase activity"
PN	WO200134785-A1.	
XX		
PD	17-MAY-2001.	
XX		
PF	10-NOV-2000; 2000WO-JP07917.	
XX		
PR	11-NOV-1999; 99JP-0321740.	
PR	16-MAY-2000; 2000JP-0144020.	
XX		
PA	(YAMA) YAMANOUCHI PHARM CO LTD.	
PA	(KAZU-) KAZUSA DNA RES INST.	
XX		
PI	Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;	
XX		
DR	WPI; 2001-343602/36.	
DR	P-PSDB; AAG62299.	
XX		
PT	Metalloprotease with aggrecanase activity for treating joint diseases	
PT	especially osteoarthritis	
XX		
PS	Example 2; Page 60-61; 85pp; Japanese.	
XX		
CC	This invention relates to a metalloprotease with aggrecanase activity.	
CC	The invention includes protein and DNA sequences of the metalloprotease,	
CC	vectors containing the DNA, host cells transformed by the vectors, and	
CC	antibodies directed against the metalloprotease. The antibodies, protein	
CC	and DNA sequences can be used in the treatment and prevention of joint	
CC	diseases, particularly osteoarthritis. The treatment may result in	
CC	osteopathic and antiarthritic activity. The present sequence represents	
CC	cDNA encoding the metalloprotease termed MDTs6.	
XX		
SQ	Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;	
	Query Match 94.5%; Score 2695; DB 22; Length 2853;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 2795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	57 TGAGCCAGAGCGGAGGTAGTCTGTTCCATCCGACTGGACCCCGACATTAAACGGCCCGG 116	
Db		
QY	57 TGAGCCAGAGCGGAGGTAGTCTGTTCCATCCGACTGGACCCCGACATTAAACGGCCCGG 116	
Db		
QY	117 CTACTACTGCGGGGTCGCCGAGACTCCGGGATCAGGACTCATTTTTCAGATCAGAC 176	
Db		
QY	117 CTACTACTGCGGGGTCGCCGAGACTCCGGGATCAGGACTCATTTTTCAGATCAGAC 176	
Db		
QY	177 ATTTTCAGGAGACTTTTACCTACACTGACCGCGGATGCTCAGTCTTGCTCCCGCCTT 236	
Db		
QY	177 ATTTTCAGGAGACTTTTACCTACACTGACCGCGGATGCTCAGTCTTGCTCCCGCCTT 236	
Db		
QY	237 CTCACCTGACATCTGGGCGTCCCTCCAGGGGCTACCGGGGCTCTTCAGACCTTCGG 296	
Db		
QY	237 CTCACCTGACATCTGGGCGTCCCTCCAGGGGCTACCGGGGCTCTTCAGACCTTCGG 296	
Db		
QY	297 ACGCTGCTTCTAATCTGGGACGCTGAACCGCGAGCGGACTCGTTCGCTGCTGAGCCT 356	
Db		

Db	297 ACGCTGCTTCTAATCTGGGACGCTGAACCGCGAGCGGACTCGTTCGCTGCTGTAGACCT 356	
QY	357 GTGCGGGGGCTCCGGAGAGCTTTGGCTACCGAGGCGCGAGTATGTCTATTAGCCCGCT 416	
Db	357 GTGCGGGGGCTCCGGAGAGCTTTGGCTACCGAGGCGCGAGTATGTCTATTAGCCCGCT 416	
QY	417 GCCCAATGCTAGCGCGCGCGGCGGAGCGCAACAGCGAGGCGCACACTTCTCCAGCG 476	
Db	417 GCCCAATGCTAGCGCGCGCGGCGGAGCGCAACAGCGAGGCGCACACTTCTCCAGCG 476	
QY	477 CCGGGGTGTTCCGGGGGGCTTCCGAGACCCACCTCTCGCTCGGGGTGCGCTCGG 536	
Db	477 CCGGGGTGTTCCGGGGGGCTTCCGAGACCCACCTCTCGCTCGGGGTGCGCTCGG 536	
QY	537 CTGGAACCCCGCCATCTACGGGCCCTTGACCCCTTACAAGCCGCGGCGGCGCTTCGG 596	
Db	537 CTGGAACCCCGCCATCTACGGGCCCTTGACCCCTTACAAGCCGCGGCGGCGCTTCGG 596	
QY	597 GGAGAGTCTAGCGGGGCGAGTCTGGGCGCGCCCAAGCGTTTCTGCTCTATCCCGGGTA 656	
Db	597 GGAGAGTCTAGCGGGGCGAGTCTGGGCGCGCCCAAGCGTTTCTGCTCTATCCCGGGTA 656	
QY	657 CGTGAGAGCGCTGGTGTGCGGACGAGTCAATGGTCAAGTTCCACGCGCGGACCTGGA 716	
Db	657 CGTGAGAGCGCTGGTGTGCGGACGAGTCAATGGTCAAGTTCCACGCGCGGACCTGGA 716	
QY	717 ACATTATCTGCTGACGCTGTGGCAACGGCGGCGACTCTACCGCCATCCCAAGTCT 776	
Db	717 ACATTATCTGCTGACGCTGTGGCAACGGCGGCGACTCTACCGCCATCCCAAGTCT 776	
QY	777 CAACCCCATCAACATCGTTGTGTCGAAGTCTCTTCTAGAGATCGTGACTCGGGCC 836	
Db	777 CAACCCCATCAACATCGTTGTGTCGAAGTCTCTTCTAGAGATCGTGACTCGGGCC 836	
QY	837 CAAGTCTACCGGCAATGCGGCCCTGACGCTGCGCAACTTCTTGCTGCGCAAGAAGCT 896	
Db	837 CAAGTCTACCGGCAATGCGGCCCTGACGCTGCGCAACTTCTTGCTGCGCAAGAAGCT 896	
QY	897 GAACAAAGTGAAGTGAACACACCCCGAGTACTGGGACACTGCCATCTTCCACAGCA 956	
Db	897 GAACAAAGTGAAGTGAACACACCCCGAGTACTGGGACACTGCCATCTTCCACAGCA 956	
QY	957 GGACCTGTGTGGAGCACCACCTGTGACACCTGGCATGGCTGATGTGGGTACCATGTG 1016	
Db	957 GGACCTGTGTGGAGCACCACCTGTGACACCTGGCATGGCTGATGTGGGTACCATGTG 1016	
QY	1017 TGACCCCAAGAGAAGTGTCTGTCTATTGAGGACGATGGGCTTCCATCAGCCTTCAAC 1076	
Db	1017 TGACCCCAAGAGAAGTGTCTGTCTATTGAGGACGATGGGCTTCCATCAGCCTTCAAC 1076	
QY	1077 TGCCACAGAGCTGGGCCAGGTGTTCAACATGCCCATGACATGTGAAGTCTGTGAGCA 1136	
Db	1077 TGCCACAGAGCTGGGCCAGGTGTTCAACATGCCCATGACATGTGAAGTCTGTGAGCA 1136	
QY	1137 GGTGTTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACG 1196	
Db	1137 GGTGTTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACG 1196	
QY	1197 TGCCAAACCCCTGGTTCAGGCTGCGAGTGTGCCATCATACCGACTTCTTGGACAGCGCA 1256	
Db	1197 TGCCAAACCCCTGGTTCAGGCTGCGAGTGTGCCATCATACCGACTTCTTGGACAGCGCA 1256	
QY	1257 CGGTGACTGCTTCCCTGGACCAACCCAGAGCCCATCTCCCTGCCAGGAGATCTGCCGG 1316	
Db	1257 CGGTGACTGCTTCCCTGGACCAACCCAGAGCCCATCTCCCTGCCAGGAGATCTGCCGG 1316	
QY	1317 CGCCAGCTACACCTTGAGCCAGCAGTGTGGCTTTTGGCGTGGGCTCCCAAGCCCTG 1376	
Db	1317 CGCCAGCTACACCTTGAGCCAGCAGTGTGGCTTTTGGCGTGGGCTCCCAAGCCCTG 1376	
QY	1377 TCCTTACATGAGTACTGCACCAAGCTGTGTGTGACCGGGAAGGCCAAGGAGACAGATGGT 1436	
Db		

Db 1377 TCCTTACATGACTACTGCACCAAGCTGTGGTGCACCGGTAAGGCCAAGGACAGATGCT 1436
Qy 1437 GTGCCAGACCCGCACTTCCCTTGGCCGATGGCACCAGCTGTGGCGAGGGCAAGCTCTG 1496
Db 1437 GTGCCAGACCCGCACTTCCCTTGGCCGATGGCACCAGCTGTGGCGAGGGCAAGCTCTG 1496
Qy 1497 CCTCAAGAGGGCTGGTGGAGAGACACAACTCAACAAGCACAGGCTGATGGTTCCTG 1556
Db 1497 CCTCAAGAGGGCTGGTGGAGAGACACAACTCAACAAGCACAGGCTGATGGTTCCTG 1556
Qy 1557 GGCAATATGGATCCCTATGGCCCTGCTCGCCGACATGTGTGGGGGCTGCAGCTGGC 1616
Db 1557 GGCAATATGGATCCCTATGGCCCTGCTCGCCGACATGTGTGGGGGCTGCAGCTGGC 1616
Qy 1617 CAGAGGCACTGACCAACCCCTCCACAGGGGTAAGTACTGCGAGGAGTGTAG 1676
Db 1617 CAGAGGCACTGACCAACCCCTCCACAGGGGTAAGTACTGCGAGGAGTGTAG 1676
Qy 1677 GGTGAATATCCGATCCCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAAGAGCTT 1736
Db 1677 GGTGAATATCCGATCCCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAAGAGCTT 1736
Qy 1737 CCGGGAGAGAGTGTGAGGCTTTCAAGGCTTCAACGGCTTACACGGCTTACACGGCTTCACTCT 1796
Db 1737 CCGGGAGAGAGTGTGAGGCTTTCAAGGCTTCAACGGCTTACACGGCTTACACGGCTTCACTCT 1796
Qy 1797 CGCGTGGCATGGTGGCCCAAGTACTCCGGCTGTCTCCCGGGACAAAGTGCAGCTCAT 1856
Db 1797 CGCGTGGCATGGTGGCCCAAGTACTCCGGCTGTCTCCCGGGACAAAGTGCAGCTCAT 1856
Qy 1857 CTGCGGAGCAATGGCACTGGCTTCTATGTCTGGCACCCCAAGGTGTGACGGCAC 1916
Db 1857 CTGCGGAGCAATGGCACTGGCTTCTATGTCTGGCACCCCAAGGTGTGACGGCAC 1916
Qy 1917 GCTGTGCTCTCTGACTCCACTCCCTGCTGTCTCAAGGCAAGTGCATCAGCTGGCTG 1976
Db 1917 GCTGTGCTCTCTGACTCCACTCCCTGCTGTCTCAAGGCAAGTGCATCAGCTGGCTG 1976
Qy 1977 TGATGGGAACCTGGCTCCAAAGAGAGATTCGACAACTGTGGGTGTGGGGGAGACAA 2036
Db 1977 TGATGGGAACCTGGCTCCAAAGAGAGATTCGACAACTGTGGGTGTGGGGGAGACAA 2036
Qy 2037 TAAGAGCTGCAAGAAGTGTGACTGACTTCTCAACAGCCCATGATGGCTACAAATTCGT 2096
Db 2037 TAAGAGCTGCAAGAAGTGTGACTGACTTCTCAACAGCCCATGATGGCTACAAATTCGT 2096
Qy 2097 GGTGCCATCCCGAGCGCTCAAGCATCGACATCGCGAGCGGTTACAAAGGCT 2156
Db 2097 GGTGCCATCCCGAGCGCTCAAGCATCGACATCGCGAGCGGTTACAAAGGCT 2156
Qy 2157 GATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCTACCTGTCAACGG 2216
Db 2157 GATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCTACCTGTCAACGG 2216
Qy 2217 GCATTTCTGTGTGGGGTGGAGCGGAGCTGGTGAAGGCACTGTGCTGCCGTA 2276
Db 2217 GCATTTCTGTGTGGGGTGGAGCGGAGCTGGTGAAGGCACTGTGCTGCCGTA 2276
Qy 2277 CAGCGGCAGCGGCAGCGGTGGAGAGCTTCAGGCTTCCCGGCCCATCTGTGAGCGCT 2336
Db 2277 CAGCGGCAGCGGCAGCGGTGGAGAGCTTCAGGCTTCCCGGCCCATCTGTGAGCGCT 2336
Qy 2337 GACCTGGAGGTTCCTTCCGTGGGGAAGATGACACCGCCCGGTGCTGCTACTCTTCTA 2396
Db 2337 GACCTGGAGGTTCCTTCCGTGGGGAAGATGACACCGCCCGGTGCTGCTACTCTTCTA 2396
Qy 2397 TCTGCCCAAGAGCTCCGGAGGACAAGTCTCTATCCCAAGGACCCCGGGACCCCTC 2456
Db 2397 TCTGCCCAAGAGCTCCGGAGGACAAGTCTCTATCCCAAGGACCCCGGGACCCCTC 2456
Qy 2457 TGTCTTGCACACACGCTCTCAGCTCTTCCAAACAGGTGGAGCGCGGACGACAGGCC 2516
Db 2457 TGTCTTGCACACACGCTCTCAGCTCTTCCAAACAGGTGGAGCGCGGACGACAGGCC 2516

Qy 2517 CCCTGCACCTGGGTGGCTGGCAGCTGGGGCCGCTGCTCCGCGAGCTGCGSCAGTGGCCT 2576
Db 2517 CCCTGCACCTGGGTGGCTGGCAGCTGGGGCCGCTGCTCCGCGAGCTGCGSCAGTGGCCT 2576
Qy 2577 GCAGAAGCGGGCTGGACTGCGGGGCTCGCGGGGAGCGGACGCTGCTGCTGCTGA 2636
Db 2577 GCAGAAGCGGGCTGGACTGCGGGGCTCGCGGGGAGCGGACGCTGCTGCTGCTGA 2636
Qy 2637 TCCAGCCCATCGCCCGTGGAGACACAGCTGCGGGGAGCCCTGCCACCTGGAGCT 2696
Db 2637 TCCAGCCCATCGCCCGTGGAGACACAGCTGCGGGGAGCCCTGCCACCTGGAGCT 2696
Qy 2697 CAGCGCTTGGTCAACCTCTCTCAAGAGCTGCGGCCGGGATTTTCAGAGGCGCTCACTCAA 2756
Db 2697 CAGCGCTTGGTCAACCTCTCTCAAGAGCTGCGGCCGGGATTTTCAGAGGCGCTCACTCAA 2756
Qy 2757 GTGTGTGGGCGACGAGCGGCTGCTGGCCGGGAGCGGACGCTGCAACTTGCACCGCAAGCC 2816
Db 2757 GTGTGTGGGCGACGAGCGGCTGCTGGCCGGGAGCGGACGCTGCAACTTGCACCGCAAGCC 2816
Qy 2817 CCAGGAGCTGGACTTCTGCTGCTGAGGCGCTGCTGA 2853
Db 2817 CCAGGAGCTGGACTTCTGCTGCTGAGGCGCTGCTGA 2853

RESULT 4

ABK12894

ID ABK12894 standard; cDNA; 2930 BP.

XX ABK12894;

XX 09-APR-2002 (first entry)

XX Human protease PRIS-11 cDNA sequence.

XX Human; protease; PRIS; gastrointestinal; Crohn's disease; cancer;
XX cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
XX inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
XX cell proliferative disorder; developmental disorder; epilepsy;
XX Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
XX reproductive disorder; endometriosis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 75..2930

XX /*tag= a

XX /partial

XX /product= "Human protease PRIS-11"

XX /note= "This sequence lacks a stop codon"

XX WO200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US19178.

XX 16-JUN-2000; 2000US-212336P.

XX 22-JUN-2000; 2000US-213955P.

XX 29-JUN-2000; 2000US-215396P.

XX 07-JUL-2000; 2000US-216821P.

XX 14-JUL-2000; 2000US-218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;

XX Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;

XX Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;

XX Azimzal Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;

XX Kaillick DA;

XX WPI; 2002-090437/12.

Pror Ant
copy
with sequence

P-PSDB; AAU74751.

Twenty one human proteases (referred to as PRPS-1 to PRPS-21), useful in the diagnosis, treatment and prevention of gastrointestinal (e.g. gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative (e.g. cancer) disorders -

Claim 5; Page 168; 177pp; English.

The present invention relates to twenty one new human proteases, referred to as PRPS-1 to PRPS-21. The PRPS polynucleotides and polypeptides of the invention are useful in the diagnosis, treatment and prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial infarction, autoimmune/inflammatory e.g. acquired immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g. cancer, developmental e.g. Duchenne and Becker muscular dystrophy, epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's disease and reproductive e.g. infertility and endometriosis disorders. Numerous other examples of each disorder are given in the specification. The present nucleic acid sequence encodes the human protease PRPS-11 protein of the invention.

XX Sequence 2930 BP; 529 A; 958 C; 931 G; 512 T; 0 other;

Query Match 77.0%; Score 2196; DB 24; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

DR	XX	661	GAGACGCTGGTGGTCCGAGAGTCAATGTCAGTTCCACGGCGGACCTGGAAACAT	720
PT	XX	735	GAGACGCTGGTGGTCCGAGAGTCAATGTCAGTTCCACGGCGGACCTGGAAACAT	794
PT	XX	721	TATCTGCTGACCTGCTGGCAACGGCGGAGCTCTACCGCCATCCAGCATCTTAAAC	780
PT	XX	795	TATCTGCTGACCTGCTGGCAACGGCGGAGCTCTACCGCCATCCAGCATCTTAAAC	854
PS	XX	781	CCATCAACATCGTTGCTCAAGTGTCTTCTTACAGATCGTACTCGGCGGCAAG	840
CC	XX	855	CCATCAACATCGTTGCTCAAGTGTCTTCTTACAGATCGTACTCGGCGGCAAG	914
CC	XX	841	GTACCGGCAATGGCGGCTGACCGTGGCGAATCTCTGTGCTGGCAGAGAACTTAAAC	900
CC	XX	915	GTACCGGCAATGGCGGCTGACCGTGGCGAATCTCTGTGCTGGCAGAGAACTTAAAC	974
CC	XX	901	AAAGTGAAGTGAACAGCAACCCGAGTACTGGGACACTGCATCTCTTCAACGAGAGAC	960
CC	XX	975	AAAGTGAAGTGAACAGCAACCCGAGTACTGGGACACTGCATCTCTTCAACGAGAGAC	1034
CC	XX	961	CTGTGTGGAGCACCACCTGTGACACCTGGGCAATGGCTGATGTGGGTACCATGTGAC	1020
CC	XX	1035	CTGTGTGGAGCACCACCTGTGACACCTGGGCAATGGCTGATGTGGGTACCATGTGAC	1094
CC	XX	1021	CCCAAGAGAGTGTCTGTCTTATTCAGGACGATGGGCTTCCATCAGCCTTCAACCAAGCC	1080
CC	XX	1095	CCCAAGAGAGTGTCTGTCTTATTCAGGACGATGGGCTTCCATCAGCCTTCAACCAAGCC	1154
CC	XX	1081	CACGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAGTCTGTGAGAGGTG	1140
CC	XX	1155	CACGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAGTCTGTGAGAGGTG	1214
CC	XX	1141	TTTGGGAAGCTCCGAGCAACCAACATGATGTCGCCGACCTCATCCAGATGACCTGACC	1200
CC	XX	1215	TTTGGGAAGCTCCGAGCAACCAACATGATGTCGCCGACCTCATCCAGATGACCTGACC	1274
CC	XX	1201	AACCCCTGTGAGCTGCTGCCATCATCAGCAGCTTCCCTGCGGAGGATCTCCGCGGCGCC	1334
CC	XX	1275	AACCCCTGTGAGCTGCTGCCATCATCAGCAGCTTCCCTGCGGAGGATCTCCGCGGCGCC	1394
CC	XX	1261	GACTGCTCTCGAGCAACCCAGCAACCTCATCTCCCTGCCGAGGATCTCCGCGGCGCC	1380
CC	XX	1335	GACTGCTCTCGAGCAACCCAGCAACCTCATCTCCCTGCCGAGGATCTCCGCGGCGCC	1394
CC	XX	1321	AGTACACACCTGAGCGAGCTGCGAGCTGGCTTTTGGCGTCCAAAGCCTTCTCTCT	1380
CC	XX	1395	AGTACACACCTGAGCGAGCTGCGAGCTGGCTTTTGGCGTCCAAAGCCTTCTCTCTCT	1454
CC	XX	1381	TACATGAGTACTGACCAAGCTGTGTGACCCGAGGAGGCAAGGACAGATGTTGTC	1440
CC	XX	1455	TACATGAGTACTGACCAAGCTGTGTGACCCGAGGAGGCAAGGACAGATGTTGTC	1514
CC	XX	1441	CAGACCCGCACTTCCCTTGGGCGGATGCGACAGCTGTGGCGAGGCAAGCTTCTGCTC	1500
CC	XX	1515	CAGACCCGCACTTCCCTTGGGCGGATGCGACAGCTGTGGCGAGGCAAGCTTCTGCTC	1574
CC	XX	1501	AAAGGGGCTGCTGGAGAGACACAACTTCAACAGCACAGGTGGATGTTCTTGGGCC	1560
CC	XX	1575	AAAGGGGCTGCTGGAGAGACACAACTTCAACAGCACAGGTGGATGTTCTTGGGCC	1634
CC	XX	1561	AAATGGATTCCTTATGGCCCTTCTCGGCGACATGTGTGGGGGCTGACGTGCGCAGG	1620
CC	XX	1635	AAATGGATTCCTTATGGCCCTTCTCGGCGACATGTGTGGGGGCTGACGTGCGCAGG	1694
CC	XX	1621	AGGAGTGCACCAACCCCACTTCCCAACGGGGGCAAGTACTGGGAGGAGTGAAGGTG	1680
CC	XX	1695	AGGAGTGCACCAACCCCACTTCCCAACGGGGGCAAGTACTGGGAGGAGTGAAGGTG	1754
CC	XX	1681	AAATCCGATCTCTCAATCTGGAGCCCTGCCCACTCAGCTCGGAAAGAGCTTCCGG	1740
CC	XX	1755	AAATCCGATCTCTCAATCTGGAGCCCTGCCCACTCAGCTCGGAAAGAGCTTCCGG	1814
CC	XX	1741	GAGGAGCTGTGAGGCTTTTCAACGGCTTACACCAACAGCAACCAACCGGCTCAGCTC	1800

Db 1815 GAGGACAGTGTGAGGCTTCAACGGCTACACACAGCACCAGCGCTCACTCTCGCC 1874
Qy 1801 GTGGATGGGTCACAGTACTCGGCGGTGTCTCCCGGGACACAGTCAAGTCACTCTGC 1860
Db 1875 GTGGATGGGTCACAGTACTCGGCGGTGTCTCCCGGGACACAGTCAAGTCACTCTGC 1934
Qy 1861 CGAGCAATGGCACTGCTACTTCTATGTGTGTGGACCCAA--GGTGTGGACGGCAGC 1917
Db 1935 CGAGCAATGGCACTGCTACTTCTATGTGTGTGGACCCAA--GGTGTGGACGGCAGC 1994
Qy 1918 CTGTGCTCTCTGACTCCACTCGCTGTGTGTCCAAAGCAAGTGCATCAAGGTGGCTGT 1977
Db 1995 CTGTGCTCTCTGACTCCACTCGCTGTGTGTCCAAAGCAAGTGCATCAAGGTGGCTGT 2054
Qy 1978 GATGGACACTGGGCTCAAGAAAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAT 2037
Db 2055 GATGGAACTGGGCTCAAGAAAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAT 2114
Qy 2038 AAGAGCTCAAGAAAGTGTGACTGCTCTTCCAAAGCCCATGTGCTATCAATTTCTGTG 2097
Db 2115 AAGAGCTCAAGAAAGTGTGACTGCTCTTCCAAAGCCCATGTGCTATCAATTTCTGTG 2174
Qy 2098 GTGGCAATCCCGAGGCGCTCAAGATCGACATCGCCAGCGCGGTACAAAGGCGTG 2157
Db 2175 GTGGCAATCCCGAGGCGCTCAAGATCGACATCGCCAGCGCGGTACAAAGGCGTG 2234
Qy 2158 ATCGGGGATGACACTACTGCTGTCTGAAGAACAGCAAGTACTGCTCAACGGG 2217
Db 2235 ATCGGGGATGACACTACTGCTGTCTGAAGAACAGCAAGTACTGCTCAACGGG 2294
Qy 2218 CATTTCTGTGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCGTGTCTCGGTAC 2277
Db 2295 CATTTCTGTGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCGTGTCTCGGTAC 2354
Qy 2278 ACCGCGAGGCGACAGCGTGGAGCGTGTGAGGCTTCCCGGCCCATCTCGAGCGCGTG 2337
Db 2355 ACCGCGAGGCGACAGCGTGGAGCGTGTGAGGCTTCCCGGCCCATCTCGAGCGCGTG 2414
Qy 2338 ACCGTGGAGGCTCTCTCGTGGGGAAGATGACACCGCCCGCGGTCTCTCTAT 2397
Db 2415 ACCGTGGAGGCTCTCTCGTGGGGAAGATGACACCGCCCGCGGTCTCTCTAT 2474
Qy 2398 CTGCCCAAGAGCTCGGAGGACAACTCTCTATCCC 2436
Db 2475 CTGCCCAAGAGCTCGGAGGACAACTCTCTATCCC 2513

RESULT 5
AAS97182
ID AAS97182 standard; cDNA; 2937 BP.
AC AAS97182;
XX AAS97182;
DT 26-FEB-2002 (first entry)
XX Human metalloprotease partial DNA sequence #11.
XX Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
KW vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;
KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
KW immune-related disease; cardiovascular disease; neuronal disease;
KW migraine; sexual dysfunction; mood disorder; attention disorder;
KW cognition disorder; hypertension; hypertention; psychotic disorder;
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.
XX Homo sapiens.
XX W0200183782-A2.
XX

PD 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US14431.
XX 04-MAY-2000; 2000US-201879P.
XX (SUGEN-) SUGEN INC.
PA Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Payne V;
XX WPI; 2002-041502/05.
DR P-PSDB; AAU72899.
XX Novel protease polypeptide useful for screening for substances that may
PT be used to treat, e.g., cancers, immune-related diseases,
PT cardiovascular disease, migraine, pain, psychotic and inflammatory
PT disorders -
XX Claim 30; Figure 1P; 232pp; English.
PS The invention relates to an isolated, enriched, or purified protease
XX polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
CC screen for substances (S) that may modulate its activity. Administering
CC S (which modulates protease activity in vitro) may be used to treat a
CC disease or disorder selected from cancers (e.g., of tissues, of blood or
CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
CC brain, ovarian, bladder or kidney), immune-related diseases and
CC disorders, cardiovascular disease, brain or neuronal-associated diseases
CC (e.g., central or peripheral nervous system diseases, migraine, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders and dyskinesias), metabolic disorders and inflammatory
CC disorders. (I) may also be useful as a diagnostic tool for a disease or
CC disorder such as those above. AAS97159-AAS97195 represent human
CC protease coding sequences and primers of the invention.
XX Sequence 2937 BP; 531 A; 958 C; 936 G; 512 T; 0 other;
SQ Query Match 77.0%; Score 2196; DB 24; Length 2937;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
Qy 1 ATGCTTCTGTGGGATCTTAAACCTTGGCTTTCGCGGGGCAACCGCTGAGGCTCTAG 60
Db 76 ATGCTTCTGTGGGATCTTAAACCTTGGCTTTCGCGGGGCAACCGCTGAGGCTCTAG 135
Qy 61 CCAGAGCGGAGGTAGTGTCTTCCATCCGACTCGGACCTTAACGCGCCGCTAC 120
Db 136 CCAGAGCGGAGGTAGTGTCTTCCATCCGACTCGGACCTTAACGCGCCGCTAC 195
Qy 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAAGCAIT 180
Db 196 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAAGCAIT 255
Qy 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240
Db 256 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 315
Qy 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGCTTTCAGACCTTGGCAGCG 300
Db 316 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGCTTTCAGACCTTGGCAGCG 375
Qy 301 TGCTTCTATTCTGGGACGTGAACCGGAGCTCGTTCTGCTGCTGTGAGCCTGTGC 360
Db 376 TGCTTCTATTCTGGGACGTGAACCGGAGCTCGTTCTGCTGCTGTGAGCCTGTGC 435
Qy 361 GGGGGGCTCCCGAGGCTTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCGCTGCC 420
Db 436 GGGGGGCTCCCGAGGCTTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCGCTGCC 495
Qy 421 AATGCTAGCGCGCGCGGCGGCGGCAACAGCCAGGCGGCGACCTTCTCCAGCGCGCG 480

Db 496 AATGCTAGCGCGCGCGGCGAGCGCAACAGCCAGGCGGCACACCTTCTCCAGCGCGG 555
QY 481 GGTGTTCCGGGCGGCGCTTCGGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG 540
Db 556 GGTGTTCCGGGCGGCGCTTCGGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG 615
QY 541 AACCCGCGCATCTACGGGCGCTTGAGCCCTTACAAAGCGCGCGGCGGGCTTCGGGGAG 600
Db 616 AACCCGCGCATCTACGGGCGCTTGAGCCCTTACAAAGCGCGCGGCGGGCTTCGGGGAG 675
QY 601 AGTCGTAGCGCGCGAGCTGTGGGCGCGCAAGGTTTCGTCTATCCCGGTACGTG 660
Db 676 AGTCGTAGCGCGCGAGCTGTGGGCGCGCAAGGTTTCGTCTATCCCGGTACGTG 735
QY 661 GAGAGCTGGTGTGCGGAGCAGGTCAATGGTCAAGTTCACGCGCGCGGACCTTGGAAAT 720
Db 736 GAGAGCTGGTGTGCGGAGCAGGTCAATGGTCAAGTTCACGCGCGCGGACCTTGGAAAT 795
QY 721 TATCTGCTACGCTGCTGCAACGCGCGGCGGACTCTACCGCCATCCCGATCCTCAAC 780
Db 796 TATCTGCTACGCTGCTGCAACGCGCGGCGGACTCTACCGCCATCCCGATCCTCAAC 855
QY 781 CCATCAACATCGTTGTGCTGAAGTGTCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
Db 856 CCATCAACATCGTTGTGCTGAAGTGTCTTCTTAGAGATCGTGACTCCGGGCCCAAG 915
QY 841 GTACCGGCAANTGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAACTGAAC 900
Db 916 GTACCGGCAANTGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAACTGAAC 975
QY 901 AAAGTGAGTGACAAGCACCCCGAGTACTGGGACACTGCGCATCTCTTCCACGAGCAGGAC 960
Db 976 AAAGTGAGTGACAAGCACCCCGAGTACTGGGACACTGCGCATCTCTTCCACGAGCAGGAC 1035
QY 961 CTGTGTGGAGCACCACTGTGACACCTTGGCATGGCTATGTGGGTACCATGTGTGAC 1020
Db 1036 CTGTGTGGAGCACCACTGTGACACCTTGGCATGGCTATGTGGGTACCATGTGTGAC 1095
QY 1021 CCCAAGAGAAGTGTCTGTCAATGAGGACGATGGGCTTCCATGAGCCTTCACCACTGCC 1080
Db 1096 CCCAAGAGAAGTGTCTGTCAATGAGGACGATGGGCTTCCATGAGCCTTCACCACTGCC 1155
QY 1081 CAGGAGCTGGGCGAGCTTCAACATGCCCCATGACATGTGAAGTCTGTGAGGAGGTG 1140
Db 1156 CAGGAGCTGGGCGAGCTTCAACATGCCCCATGACATGTGAAGTCTGTGAGGAGGTG 1215
QY 1141 TTGTGGAAGCTCCGAGCAACACATGATGCCCGACCTTCATCCAGATCGACCGTGGC 1200
Db 1216 TTGTGGAAGCTCCGAGCAACACATGATGCCCGACCTTCATCCAGATCGACCGTGGC 1275
QY 1201 AACCCCTGGTACGCTGCAAGTGTGCTGATCATCACCGACTTCTGGAGCAGCGGACGGT 1260
Db 1276 AACCCCTGGTACGCTGCAAGTGTGCTGATCATCACCGACTTCTGGAGCAGCGGACGGT 1335
QY 1261 GACTGCTCTCTGACCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
Db 1336 GACTGCTCTCTGACCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1395
QY 1321 AGCTACACCTTGAGCAGCAGTGGAGCTGGCTTTTGGGCTGGGCTCCAAAGCCCTGTCT 1380
Db 1396 AGCTACACCTTGAGCAGCAGTGGAGCTGGCTTTTGGGCTGGGCTCCAAAGCCCTGTCT 1455
QY 1381 TACATCAGTACTGCAACCAAGTGTGTGTCACCGGGAAGGCAAGGACATGTTGTGC 1440
Db 1456 TACATCAGTACTGCAACCAAGTGTGTGTCACCGGGAAGGCAAGGACATGTTGTGC 1515
QY 1441 CAGACCGGCACTTCCCTGGGCGGATGGCAACAGCTGTGGCGAGGCAAGCTCTGCCTC 1500
Db 1516 CAGACCGGCACTTCCCTGGGCGGATGGCAACAGCTGTGGCGAGGCAAGCTCTGCCTC 1575
QY 1501 AAAGGGGCGCTGCTGGAGAGACACACCTCAACAGACGACAGGCTGGATGGTCTCTGGGCC 1560
Db 1576 AAAGGGGCGCTGCTGGAGAGACACACCTCAACAGACGACAGGCTGGATGGTCTCTGGGCC 1635

QY 1561 AAATGGATCCCTATGTCCTGCTCGGCACATCTGTGTGGGGCGTGCAGCTGGCCAGG 1620
Db 1636 AAATGGATCCCTATGTCCTGCTCGGCACATCTGTGTGGGGCGTGCAGCTGGCCAGG 1695
QY 1621 AGGCAAGTGCACCAACCCCGCTGCGCAACCGGGGCAAGTACTCGAGGAGGTGAGGGTG 1680
Db 1696 AGGCAAGTGCACCAACCCCGCTGCGCAACCGGGGCAAGTACTCGAGGAGGTGAGGGTG 1755
QY 1681 AAATACCGATCTGCAATCTGAGCCCTGCCCGCAGCTCAGCCTCCGAAAGAGCTTCCCG 1740
Db 1756 AAATACCGATCTGCAATCTGAGCCCTGCCCGCAGCTCAGCCTCCGAAAGAGCTTCCCG 1815
QY 1741 GAGGAGAGTGTGAGGCTTTCAACGGCTTAAACACAGCAGCAACCGGCTCACTCTCCCG 1800
Db 1816 GAGGAGAGTGTGAGGCTTTCAACGGCTTAAACACAGCAGCAACCGGCTCACTCTCCCG 1875
QY 1801 GTGGCATGGTGCACCAAGTACTCCGGCGTGTCTCCGGGCAAGTGCAGCTCATCTGC 1860
Db 1876 GTGGCATGGTGCACCAAGTACTCCGGCGTGTCTCCGGGCAAGTGCAGCTCATCTGC 1935
QY 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGACCCAA ---GTTGTGGAGCGGACG 1917
Db 1936 CGAGCAATGGCACTGGCTTCTATGTGTGGACCCAAAGTGTGTGGAGCGGACG 1995
QY 1918 CTGTGCTCTCTGACTCCGCTGTGTCCAAAGCAAGTGCATCAAGCCTGGCTGT 1977
Db 1996 CTGTGCTCTCTGACTCCGCTGTGTCCAAAGCAAGTGCATCAAGCCTGGCTGT 2055
QY 1978 GATGGAACCTGGGCTCCAAAGAGATTCGACAGTGTGGGTGTGTGGGGGAGACAAT 2037
Db 2056 GATGGAACCTGGGCTCCAAAGAGATTCGACAGTGTGGGTGTGTGGGGGAGACAAT 2115
QY 2038 AAGGCTCAAGAGTGTGACTGTCTTCCAGGCGCATGCTGCTACAATTTCTGTG 2097
Db 2116 AAGGCTCAAGAGTGTGACTGTCTTCCAGGCGCATGCTGCTACAATTTCTGTG 2175
QY 2098 GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGCGGTTACAAAGGCGTG 2157
Db 2176 GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGCGGTTACAAAGGCGTG 2235
QY 2158 ATCGGGATGCAACTACTGTGCTCTGAAGACGCAAGCAAGTACTCTCAAGCGG 2217
Db 2236 ATCGGGATGCAACTACTGTGCTCTGAAGACGCAAGCAAGTACTCTCAAGCGG 2295
QY 2218 CATTTCTGTGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTCGGTAC 2277
Db 2296 CATTTCTGTGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTCGGTAC 2355
QY 2278 AGCGGCACGCGACGCGGTGGAGCGCTTCCCGCCATCGAGGCTTCCCGCCATCGTGGAGCGGCTG 2337
Db 2356 AGCGGCACGCGCGAGCGGTGGAGCGCTTCCCGCCATCGAGGCTTCCCGCCATCGTGGAGCGGCTG 2415
QY 2338 ACCTGTGAGTCTCTCTCCGTGGGGAAGTGAACCCCGCGGCTCGCTACTCTCTCTAT 2397
Db 2416 ACCTGTGAGTCTCTCTCCGTGGGGAAGTGAACCCCGCGGCTCGCTACTCTCTCTAT 2475
QY 2398 CTGCCCCAAGAGCTCTGGGAGGACAAAGTCTCTATCCC 2436
Db 2476 CTGCCCCAAGAGCTCTGGGAGGACAAAGTCTCTATCCC 2514

RESULT 6

AAD35570

ID AAD35570 standard; cDNA; 1104 BP.

XX AAD35570;

XX 26-JUL-2002 (first entry)

XX Human protease cDNA #3.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;

KW

high blood pressure; arthritis; connective tissue disorder; infertility;
gene therapy; enzyme; gene; ss.
Homo sapiens.

Key Location/Qualifiers
CDS 1..1104
/tag= a
/product= "Human protease #3"

WO200226949-A2.

04-APR-2002.

27-SEP-2001; 2001WO-US030350.

29-SEP-2000; 2000US-236689P.

(LEXI-) LEXICON GENETICS INC.

Friddle CJ, Hilbun E;

WPI; 2002-372123/40.

P-PSDB; AAE22542.

Novel nucleic acid encoding a human protease, useful as a hybridization probe for screening libraries and assessing gene expression patterns -
Disclosure; Page 38-39; 41pp; English.

The present sequence is a cDNA encoding novel human protein (NHP), human protease. NHPs share structural similarity with animal proteases particularly zinc metalloproteases. Sequences of the invention are useful in therapeutic, diagnostic and pharmacogenomic applications. NHP polynucleotides are used as hybridisation probes for screening libraries and assessing gene expression patterns. They can also be used for treating related biological disorders such as obesity, high blood pressure, arthritis, connective tissue disorders and infertility. They are also used in gene therapy.

Sequence 1104 BP; 189 A; 357 C; 345 G; 213 T; 0 other;

Query Match 38.2%; Score 1091; DB 24; Length 1104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGGTAGTGTTCCTCCATCCGACTGGAGCCGACATTAACGGCCGCGGTAC 120
DB 61 CCAGAGCGGGAGGTAGTGTTCCTCCATCCGACTGGAGCCGACATTAACGGCCGCGGTAC 120
QY 121 TACTGGGGGCTCCGGAGGACTCCGGGATCAGGACTCATTTTTCAGATCAGACATTT 180
DB 121 TACTGGGGGCTCCGGAGGACTCCGGGATCAGGACTCATTTTTCAGATCAGACATTT 180
QY 181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTCTTCTGGCTCCGCGCTTCC 240
DB 181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTCTTCTGGCTCCGCGCTTCC 240
QY 241 ACTGAGCATCTGGGGTCCCGCTCCAGGGGCTCAGGGGCTCTTCAGACCTCGGACGC 300
DB 241 ACTGAGCATCTGGGGTCCCGCTCCAGGGGCTCAGGGGCTCTTCAGACCTCGGACGC 300
QY 301 TCGTCTTATCTGGGAGCTGAACCGCGGACTCGTTCGCTGTGTGAGCTCTGTC 360
DB 301 TCGTCTTATCTGGGAGCTGAACCGCGGACTCGTTCGCTGTGTGAGCTCTGTC 360
QY 361 GGGGGGCTCCGCGGAGCCTTTGGTACCAGGCGCGGAGTATGTCATTAGCCCGCTGCC 420
DB 361 GGGGGGCTCCGCGGAGCCTTTGGTACCAGGCGCGGAGTATGTCATTAGCCCGCTGCC 420

QY 421 AATGCTAGCGCGCGGGCGGCGAGCGCAACAGCGCGCGCACACCTTCTCCAGCGCGG 480
DB 421 AATGCTAGCGCGCGGGCGGCGAGCGCAACAGCGCGCGCACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGGGGGGCTTCCGGAGAGCCCACTCTCGCTCGGGGTGGCTCGGGCTGG 540
DB 481 GGTGTTCCGGGGGGGCTTCCGGAGAGCCCACTCTCGCTCGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTACGGGGCTTACAGCCCTTACAGCGCGCGGGGGGCTTCGGGGAG 600
DB 541 AACCCCGCATCTACGGGGCTTACAGCCCTTACAGCGCGCGGGGGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGGAGGTCTGGGCGGCCAAGCGTTTCTGTCTATCCCGCGGTAGCTG 660
DB 601 AGTCGTAGCGCGCGGAGGTCTGGGCGGCCAAGCGTTTCTGTCTATCCCGCGGTAGCTG 660
QY 661 GAGAGCTGTGGTTCGCGGAGGAGTCAATGTCAGTTCACGCGCGGAGCTTGAACAT 720
DB 661 GAGAGCTGTGGTTCGCGGAGGAGTCAATGTCAGTTCACGCGCGGAGCTTGAACAT 720
QY 721 TATCTGCTGAGCGTCTGGCAACGCGCGGAGCTTACCGCCATCCAGCATCTCAAC 780
DB 721 TATCTGCTGAGCGTCTGGCAACGCGCGGAGCTTACCGCCATCCAGCATCTCAAC 780
QY 781 CCCATCAACATGTTGTGGTCAAGTGTCTTCTTAGAGATCGTACTCGGCGCCCAAG 840
DB 781 CCCATCAACATGTTGTGGTCAAGTGTCTTCTTAGAGATCGTACTCGGCGCCCAAG 840
QY 841 GTCACCGCAATGCGCGCTGACGCTGCGCAACTTCTGCTGGCGAGAAAGCTCAAC 900
DB 841 GTCACCGCAATGCGCGCTGACGCTGCGCAACTTCTGCTGGCGAGAAAGCTCAAC 900
QY 901 AAAGTGAAGTGAACAGCAACCGGAGTACTGGGACATGCGCATCTTCCAGGAGGAC 960
DB 901 AAAGTGAAGTGAACAGCAACCGGAGTACTGGGACATGCGCATCTTCCAGGAGGAC 960
QY 961 CTGTGTGGAGCCACCATCTGTGACACCTGGCATGGCTGTGATGGGTACCATGTGAC 1020
DB 961 CTGTGTGGAGCCACCATCTGTGACACCTGGCATGGCTGTGATGGGTACCATGTGAC 1020
QY 1021 CCCAAGAGAAGCTGCTCTGTCATTGAGGACGATGGGCTTCCATCAGCTTCCACTGCC 1080
DB 1021 CCCAAGAGAAGCTGCTCTGTCATTGAGGACGATGGGCTTCCATCAGCTTCCACTGCC 1080
QY 1081 CACGAGCTGGG 1091
DB 1081 CACGAGCTGGG 1091
RESULT 7
AAD35568
ID AAD35568 standard; cDNA; 966 BP.
XX
AC AAD35568;
XX
DT 26-JUL-2002 (first entry)
XX
DE Human protease cDNA #1.
XX
DE Human; novel human protein; NHP; protease; biological disorder; obesity;
KW high blood pressure; arthritis; connective tissue disorder; infertility;
KW gene therapy; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..966
FT CDS /tag= a
FT /product= "Human protease #1"
XX
PN WO200226949-A2.
XX

PD	04-APR-2002.	
XX		
PF	27-SEP-2001; 2001WO-US30350.	
XX		
PR	29-SEP-2000; 2000US-236689P.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
XX		
PI	Friddle CJ, Hilbun E;	
XX		
DR	WPI: 2002-372123/40.	
XX		
DR	P-PSDB; AAE22540.	
XX		
PT	Novel nucleic acid encoding a human protease, useful as a hybridization	
XX	probe for screening libraries and assessing gene expression patterns -	
PS	Disclosure; Page 34; 41pp; English.	
XX		
CC	The present sequence is a cDNA encoding novel human protein (NHP),	
CC	human protease. NHPs share structural similarity with animal proteases	
CC	particularly zinc metalloproteases. Sequences of the invention are	
CC	useful in therapeutic, diagnostic and pharmacogenomic applications.	
CC	NHP polynucleotides are used as hybridisation probes for screening	
CC	libraries and assessing gene expression patterns. They can also be	
CC	used for treating related biological disorders such as obesity, high	
CC	blood pressure, arthritis, connective tissue disorders and infertility.	
XX	They are also used in gene therapy.	
SQ	Sequence 966 BP; 161 A; 316 C; 305 G; 184 T; 0 Other;	
Query Match 33.6%; Score 959; DB 24; Length 966;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG 60	
Db	1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG 60	
QY	61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGACCGCGACATTAACGGCGCGCTAC 120	
Db	61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGACCGCGACATTAACGGCGCGCTAC 120	
QY	121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT 180	
Db	121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT 180	
QY	181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTGGCTCCCGCTCTCC 240	
Db	181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTGGCTCCCGCTCTCC 240	
QY	241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTTTTCAGACCTGCGACGC 300	
Db	241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTTTTCAGACCTGCGACGC 300	
QY	301 TGCTTCTATTCTGGGAGGTGAAGCGCGAGCGCGGACTGCTGCTGTGCTGAGCCTGTC 360	
Db	301 TGCTTCTATTCTGGGAGGTGAAGCGCGAGCGCGGACTGCTGCTGTGCTGAGCCTGTC 360	
QY	361 GGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGCGAGTATGTCATTAGCCGCTGCC 420	
Db	361 GGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGCGAGTATGTCATTAGCCGCTGCC 420	
QY	421 AATGCTAGCGCGCGGGCGGACGCGCAACAGCGAGCGCGCACACCTTCTCCAGCGCGG 480	
Db	421 AATGCTAGCGCGCGGGCGGACGCGCAACAGCGAGCGCGCACACCTTCTCCAGCGCGG 480	
QY	481 GGTGTCCGGCGGGCTTCGGAGACCCACCTCTCGTGGGGTGGCTCGGGCTGG 540	
Db	481 GGTGTTCGGCGGGCTTCGGAGACCCACCTCTCGTGGGGTGGCTCGGGCTGG 540	
QY	541 AACCCCGCATCTTACGGGCGCTGAGCCCTTACAAGCGCGGGGGGGCTTCGGGCGAG 600	
Db	541 AACCCCGCATCTTACGGGCGCTGAGCCCTTACAAGCGCGGGGGGGCTTCGGGCGAG 600	

04-APR-2002.

27-SEP-2001; 2001WO-US30350.

29-SEP-2000; 2000US-236689P.

(LEXI-) LEXICON GENETICS INC.

Friddle CJ, Hilbun E;

WPI: 2002-372123/40.

P-PSDB; AAE22540.

Novel nucleic acid encoding a human protease, useful as a hybridization probe for screening libraries and assessing gene expression patterns -

Disclosure; Page 34; 41pp; English.

The present sequence is a cDNA encoding novel human protein (NHP), human protease. NHPs share structural similarity with animal proteases particularly zinc metalloproteases. Sequences of the invention are useful in therapeutic, diagnostic and pharmacogenomic applications. NHP polynucleotides are used as hybridisation probes for screening libraries and assessing gene expression patterns. They can also be used for treating related biological disorders such as obesity, high blood pressure, arthritis, connective tissue disorders and infertility. They are also used in gene therapy.

Sequence 966 BP; 161 A; 316 C; 305 G; 184 T; 0 Other;

Query Match 33.6%; Score 959; DB 24; Length 966;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG 60

1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG 60

61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGACCGCGACATTAACGGCGCGCTAC 120

61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGACCGCGACATTAACGGCGCGCTAC 120

121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT 180

121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT 180

181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTGGCTCCCGCTCTCC 240

181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTGGCTCCCGCTCTCC 240

241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTTTTCAGACCTGCGACGC 300

241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTTTTCAGACCTGCGACGC 300

301 TGCTTCTATTCTGGGAGGTGAAGCGCGAGCGCGGACTGCTGCTGTGCTGAGCCTGTC 360

301 TGCTTCTATTCTGGGAGGTGAAGCGCGAGCGCGGACTGCTGCTGTGCTGAGCCTGTC 360

361 GGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGCGAGTATGTCATTAGCCGCTGCC 420

361 GGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGCGAGTATGTCATTAGCCGCTGCC 420

421 AATGCTAGCGCGCGGGCGGACGCGCAACAGCGAGCGCGCACACCTTCTCCAGCGCGG 480

421 AATGCTAGCGCGCGGGCGGACGCGCAACAGCGAGCGCGCACACCTTCTCCAGCGCGG 480

481 GGTGTCCGGCGGGCTTCGGAGACCCACCTCTCGTGGGGTGGCTCGGGCTGG 540

481 GGTGTTCGGCGGGCTTCGGAGACCCACCTCTCGTGGGGTGGCTCGGGCTGG 540

541 AACCCCGCATCTTACGGGCGCTGAGCCCTTACAAGCGCGGGGGGGCTTCGGGCGAG 600

541 AACCCCGCATCTTACGGGCGCTGAGCCCTTACAAGCGCGGGGGGGCTTCGGGCGAG 600

601 AGTCGTAGCCGGCGCAGGCTCTGGCGCGCCCAAGCGTTTCTGTCTATCCCGCGTACGTG 660

601 AGTCGTAGCCGGCGCAGGCTCTGGCGCGCCCAAGCGTTTCTGTCTATCCCGCGTACGTG 660

661 GAGACGCTGGTGTGCGCGGAGCTCAATGTCAGTTCCACGCGCGGACCTGGACAT 720

661 GAGACGCTGGTGTGCGCGGAGCTCAATGTCAGTTCCACGCGCGGACCTGGACAT 720

721 TATCTGTGACGCTGTGGCAACGCGCGGCTCTTACCGCCATCCAGCATCTCTCAAC 780

721 TATCTGTGACGCTGTGGCAACGCGCGGCTCTTACCGCCATCCAGCATCTCTCAAC 780

781 CCCATCAACATCTGTGGTCAAGGCTCTCTTCTTAGAGATCTGACTCCGGGCGCAAG 840

781 CCCATCAACATCTGTGGTCAAGGCTCTCTTCTTAGAGATCTGACTCCGGGCGCAAG 840

841 GTCAACCGCAATCGGCGCTGAGCTGCGCAACTTCTGTCCCTGGCAGAGAACTGAAC 900

841 GTCAACCGCAATCGGCGCTGAGCTGCGCAACTTCTGTCCCTGGCAGAGAACTGAAC 900

901 AAAGTGAGTGACAGCAACCGCGGCTGAGTGGGACACTGCTCTTCCAGGCGAGGA 959

901 AAAGTGAGTGACAGCAACCGCGGCTGAGTGGGACACTGCTCTTCCAGGCGAGGA 959

RESULT 8

AA95831

ID AAA95831 standard; cDNA; 1143 BP.

XX AC AAA95831;

XX DT 23-FEB-2001 (first entry)

XX DE Human metalloproteinase ADAMTS-5 cDNA.

XX KW Human; ADAMTS-5; metalloproteinase; ADAM;

XX KW a disintegrin and metalloproteinase domain; thrombospondin domain;

XX KW vaccine; neutropic; neuroprotective; antiparkinsonian;

XX KW cerebroprotective; cytosolic; antiarthritis; immunosuppressive;

XX KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;

XX KW autoimmune disease; brain tumour; brain injury; ss.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 1..1143

XX FT /*tag= a

XX FT /partial

XX FT /product= "ADAMTS-5"

XX PN WO200053774-A2.

XX PD 14-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US06237.

XX PF 08-MAR-1999; 99US-0264585.

XX PR (NEUR-) NEUROCRINE BIOSCIENCES INC.

XX PA Kelner GS, Clark M, Maki RA;

XX PI WPI: 2000-594326/56.

XX DR P-PSDB; AAB21261.

XX DR Polynucleotide encoding novel members of a disintegrin,

XX PT metalloproteinase and thrombospondin domain protein family used to

XX PT prevent and treat Alzheimer's disease, cancer and autoimmune diseases -

XX PS Claim 2; Fig 22; 129pp; English.

XX CC The present sequence encodes human metalloproteinase ADAMTS-5. The


```

PR 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
PS Example 13; Page 66; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC an adaptor primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 30 BP; 8 A; 6 C; 10 G; 6 T; 0 other;

Query Match 1.1%; Score 30; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 TTACTACCTACCTGACGCCGATGCTCAGT 220
DB 30 TTTACTACCTACCTGACGCCGATGCTCAGT 1

RESULT 11
AAH41030/c
ID AAH41030 standard; DNA; 30 BP.
XX
AC AAH41030;
XX
DT 23-AUG-2001 (first entry)
DE PCR primer specific for metalloprotease DNA isolation.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; primer; adaptor; ss.
XX
OS Homo sapiens.
XX
PN WO200134785-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-JP07917.
XX
PR 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
PS Example 13; Page 66; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.

```

```

CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC an adaptor primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 30 BP; 10 A; 6 C; 6 G; 8 T; 0 other;

```

```

Query Match 1.1%; Score 30; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 TTTTTCAGATCAGCAGCATTTTCAGGAGGACT 190
DB 30 TTTTTCAGATCAGCAGCATTTTCAGGAGGACT 1

```

```

RESULT 12
AAH41017/c
ID AAH41017 standard; DNA; 41 BP.
XX
AC AAH41017;
XX
DT 23-AUG-2001 (first entry)
DE PCR primer specific for human metalloprotease DNA.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200134785-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-JP07917.
XX
PR 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
PS Example 4; Page 62; 85pp; Japanese.
XX

```

```

CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 41 BP; 5 A; 16 C; 13 G; 7 T; 0 other;

```

```

Query Match 1.1%; Score 30; DB 22; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1720 GCCTCCGAAAGAGCTTCGGGAGGAGCAG 1749

```

Db 41 GCCTCGGAAGAGCTTCGGGAGGAGCAG 12
|||||

RESULT 13

AAH41018
ID AAH41018 standard; DNA; 27 BP.

XX AC AAH41018;
XX DT 23-AUG-2001 (first entry)

XX PCR primer specific for human metalloprotease DNA SEQ ID 9.

XX Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
XX osteopathic; antiarthritic; PCR primer; ss.

XX OS Homo sapiens.

XX PN WO200134785-A1.

XX PD 17-MAY-2001.

XX PF 10-NOV-2000; 2000WO-JP07917.

XX PR 11-NOV-1999; 99JP-0321740.

XX PR 16-MAY-2000; 2000JP-0144020.

XX PA (YAMA) YAMANOUCHI PHARM CO LTD.

XX PA (KAZU-) KAZUSA DNA RES INST.

XX PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

XX DR WPI; 2001-343602/36.

XX PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis -

XX PS Example 5; Page 63; 85pp; Japanese.

XX CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.

XX SQ Sequence 27 BP; 5 A; 6 C; 11 G; 5 T; 0 other;

Query Match 0.9%; Score 27; DB 22; Length 27;

Best Local Similarity 100.0%; Pred. No. 0.044;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1534 AAGCACAGGTGGATGTTCTCTGGGCC 1560

Db 1 AAGCACAGGTGGATGTTCTCTGGGCC 27
|||||

RESULT 14

AAH41019/c

ID AAH41019 standard; DNA; 37 BP.

XX AC AAH41019;

XX DT 23-AUG-2001 (first entry)

XX PCR primer specific for human metalloprotease DNA SEQ ID 10.

XX Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
XX osteopathic; antiarthritic; PCR primer; ss.

XX OS Homo sapiens.

XX PN WO200134785-A1.

XX PD 17-MAY-2001.

XX PF 10-NOV-2000; 2000WO-JP07917.

XX PR 11-NOV-1999; 99JP-0321740.

XX PR 16-MAY-2000; 2000JP-0144020.

XX PA (YAMA) YAMANOUCHI PHARM CO LTD.

XX PA (KAZU-) KAZUSA DNA RES INST.

XX PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

XX DR WPI; 2001-343602/36.

XX PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis -

XX PS Example 5; Page 63; 85pp; Japanese.

XX CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.

XX SQ Sequence 37 BP; 7 A; 14 C; 14 G; 2 T; 0 other;

Query Match 0.9%; Score 27; DB 22; Length 37;

Best Local Similarity 100.0%; Pred. No. 0.043;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2824 CTGGACTTCTCGTCTCGAGCCGTGC 2850

Db 37 CTGGACTTCTCGTCTCGAGCCGTGC 11
|||||

RESULT 15

AAH41024/c

ID AAH41024 standard; DNA; 38 BP.

XX AC AAH41024;

XX DT 23-AUG-2001 (first entry)

XX PCR primer specific for metalloprotease related DNA.

XX Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;

XX osteopathic; antiarthritic; PCR primer; ss.

XX OS Homo sapiens.

XX PN WO200134785-A1.

XX PD 17-MAY-2001.

XX PF 10-NOV-2000; 2000WO-JP07917.

XX PR 11-NOV-1999; 99JP-0321740.

XX PR 16-MAY-2000; 2000JP-0144020.

XX PA (YAMA) YAMANOUCHI PHARM CO LTD.

XX PA (KAZU-) KAZUSA DNA RES INST.

XX PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

```
XX WPI; 2001-343602/36.
XX Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
XX Example 9; Page 65; 85pp; Japanese.
XX This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
XX Sequence 38 BP; 6 A; 13 C; 8 G; 11 T; 0 other;
SQ
Query Match 0.9%; Score 27; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2035 ATAAGAGCTGCAAGAAGTGACTGGA 2061
DB 38 ATAAGAGCTGCAAGAAGTGACTGGA 12
|||||
|||||
```

Search completed: May 16, 2003, 01:01:50
Job time : 587 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:21:31 ; Search time 29 Seconds
(without alignments)
963.854 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLLGILTLAFAGRTAGGSE.....DQCNLHRKPOHDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2486.5	48.2	967	4	US-09-130-491-2
2	2115	41.0	905	4	US-09-369-364A-9
3	1915	37.1	837	4	US-09-122-126B-2
4	1900	36.8	608	4	US-09-130-491-13
5	1854	35.9	930	4	US-09-369-364A-2
6	1848.5	35.8	1882	4	US-09-369-364A-13
7	1847.5	35.8	930	4	US-09-122-126B-15
8	1798.5	34.8	874	4	US-09-369-364A-15
9	1607	31.1	551	4	US-09-130-491-16
10	1273	24.7	481	4	US-09-130-491-8
11	1264	24.5	518	4	US-09-369-364A-22
12	1248.5	24.2	1224	4	US-09-930-872-4
13	1218	23.6	1211	4	US-09-491-522-5
14	1205.5	23.4	997	4	US-09-369-364A-7
15	1152	22.3	1081	4	US-09-369-364A-17
16	1144	22.2	1205	4	US-09-491-522-11
17	1035	20.1	859	4	US-09-369-364A-5
18	662.5	12.8	245	4	US-09-369-364A-11
19	577.5	11.2	566	4	US-09-491-522-7
20	550.5	10.7	525	4	US-09-369-364A-2h
21	441	8.5	491	4	US-09-930-872-2
22	380.5	7.4	450	4	US-09-369-364A-19
23	349	6.8	812	4	US-09-632-098-4
24	339	6.6	802	4	US-09-632-098-2
25	306	5.9	464	4	US-09-411-329C-14
26	305.5	5.9	616	4	US-09-608-790-1
27	304	5.9	462	4	US-09-411-329C-3

```

28 304 5.9 462 4 US-09-411-329C-17 Sequence 17, Appl
29 301 5.8 529 2 US-08-836-442-3 Sequence 3, Appl
30 288 6.21 4 US-09-026-001A-6 Sequence 6, Appl
31 285 5.5 855 4 US-09-813-819-2 Sequence 2, Appl
32 285 5.5 855 4 US-09-920-048-2 Sequence 2, Appl
33 281 5.4 621 4 US-09-026-001A-18 Sequence 18, Appl
34 279 5.4 769 1 US-08-243-542-4 Sequence 4, Appl
35 279 5.4 769 1 US-08-477-407-4 Sequence 4, Appl
36 279 5.4 769 1 US-08-484-355-4 Sequence 4, Appl
37 277.5 5.4 814 4 US-09-813-819-4 Sequence 4, Appl
38 277.5 5.4 814 4 US-09-920-048-4 Sequence 4, Appl
39 276 5.3 613 4 US-09-026-001A-10 Sequence 10, Appl
40 274.5 5.3 751 2 US-08-836-442-3 Sequence 3, Appl
41 263 5.1 592 4 US-09-026-001A-14 Sequence 14, Appl
42 260 5.0 670 1 US-08-243-542-3 Sequence 3, Appl
43 260 5.0 670 1 US-08-477-407-3 Sequence 3, Appl
44 260 5.0 670 1 US-08-484-355-3 Sequence 3, Appl
45 248.5 4.8 1170 1 US-08-313-288B-20 Sequence 20, Appl

```

ALIGNMENTS

RESULT 1

```

US-09-130-491-2
; Sequence 2, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; EARLIER FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-130-491-2

```

```

Query Match 48.2%; Score 2486.5; DB 4; Length 967;
Best Local Similarity 49.0%; Pred. No. 1.5e-198;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

```

```

QY 1 MLLLGILTLAFAGRTAGGSEPREVVVPIRLDPDINGRRYYWRGPEDSGDGLIFQITAF 60
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 36 LLLLAALLAVSALGRPEDEELVVP-ELE-----RAP---GHGTRRLRHAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGSSDLRRCFYSGDVNAEPDSF 114
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 82 DQQLDLRLPDSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAYSLCGGLRGATGYGAEVVISPLPNAS---APAAQRNSQGA---HLLQ---RGVPGG 165
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 137 AALLSGVNGAFYLLGEAYFIQPLPAASERLATAPEKPPAPLOFHLRLKRNUGDVG 196
QY 166 PSG-----DPTSRC-----GVASG--WNPAILLALDPYKPRRAGFSGESRRRS 207
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 197 TCGVWDEPRPTKATEDEDETEDETEDETEDETEDETEDETEDETEDETEDETEDE 247
QY 208 GRAKREVSIPRYVETLVADSEVMKFGADLEHYLLTLATLAARLYRHPISILNPINIVV 267
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 248 IRKKREVSHRYVETMLVADQSMAEFGSLKHYLLTLFSAARLYKHPISIRNSVLVV 307
QY 268 KVLRLDRSGPKVTGNAALTILNFCAWOKKLNVSKDHPYWDRTAILFTRODLGATTC 327
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 308 KILVHDEQGPVTSNAALTILNFCAWOKKLNVSKDHPYWDRTAILFTRODLGATTC 367

```


; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-122-126B-2

Query Match 37.1%; Score 1915; DB 4; Length 837;
Best Local Similarity 46.3%; Pred. No. 5,4e-151;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

```
QY 1 MLLGLTILAFAGRAGGSEPREVVVPIRLPDINGRRYYWRGPEDSGCQLFIQTAF 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 LLLLLLALSLPASLPREEIIVPEKLGSLV-----PGSGAPARLLCKLQAF 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 QEDFYHLPLDPAQFLAPAFSTHGLVPLQGLGGSDLLRRCPYSGDVNAELPDSFAVSLC 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 GETLLELEQDSGVQEGVTLVYLQAPB-LUGGAEP--GYLTGTGTLNGLHESVASLHMD 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 GG-LRGAFYGAEEYVISPLNPASAPAAQNSQGAHLQRRGVPGPSGLPTSRCCVASG 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 GGALLGLVLYRGAELHLQPLEGGTNSA--GGPGAHLRRK----SPASQGPNCNV--- 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 WNPAILRALDPYKPRAGGSESRSSRRGRKRFVSIPIRYVETLVVADESVMKPHGADLE 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 -----KAPLGSPSPRP--RAKRFASLRFVETLVVADDRKHAAPHGAGLK 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 HYLLTLLAARLYRHPSTILNPINWVVKVLLLRDRDGPVKVTGNAALTNRNFCANQKLL 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 RYLLVTMAAAKAFKHPSTRNPVSLVTRVLVLSGEGEPQVGPSSAQTITKSCAWQRL 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 NKVSKHPEYWDTAILFTQDLGGATCDTTLGMADVMTCDPKRSCSVIADGLPSAFTT 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 NTPEDSDPHEDTAILFTQDLGGVSTCDTTLGMADVMTCDPARSCAIVEDDGLQSAFTA 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 AHELGHVENMHDNVKCEEEYVGLI--RANHMSPTLIQIDRANPWSACSAAIITDFDLSG 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 AHELGHVENMHDNKPCLISLNGPLUSTSRHVNAPVMAHVDPEEPWSPCSARFITDFDLSG 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 419 HGCDLLDQSPKPISLPEDLGASTYLSQOCELAFGVSGSKPCPYM--QYCTKLWCTGKAKGQ 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 YGHCLLDKPEAPLHPVTFPGKYDADRCQLITFGPDSRHCPQLPPCAALWCSGLNHG 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 478 MYCOTRHPHWADGTCGEGEKLCLGACVERINLNKHRVD--GSAKWMDPYGPCSRTCGG 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 480 AMCOTKHPWADGTCGPAQACMGRCCLHMDQLQDFNIPQAGGWPMPGADCSRTCGG 539
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 536 VOLARRQCTNPTPANGKGYCEGVRYKRSNLEPCPSSASGKSPREOCEAFNGYNHSTN 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 540 VQFSSRDCTRPVRNGKGYCEGRTRFRSCNTEDCP--TGSALTREBOCA--YINHTD 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 RLTL---AVAWPKYSGVSPRDKCLICRANGTGYFYVYAPKVVVDGTLCSPDSTSVQVQ 652
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 596 LFKSPGPMWVPRYTGVAPOQCKLTQARALGYVYVLEPRVVDGTPCSPDSSVCVQ 655
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 653 KCIRAGCDNGLSKKREDKCVCGGDNKSKCKVTLGTLKPMHGFVVAIPAGASSIDIR 712
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 656 KCIHAGCDRIIGSKKKFKDKCMKGGGCGGCKSGSKGSKFRKFRYNNVVTIPAGATHILVR 715
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 713 ORGYKGLIGDNNYLALNKSQKYLINGHFVVSASVERDLVVKGSF--LRYSGTGTAVESLQA 771
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 716 QOQNGP--HRSIYLALKLPGSYALNGEYTLMPSTDVLPAGVSLRYSGATASETLSG 773
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 772 SRPILPTVSVLGVKMTPPRYSFYLPKPREDKSSHPKD 814
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 774 HGPLAQPLTLQVLVAGNPQDTRLYSFVFPRTPTSTPRTPQD 816
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 4

US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130.491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-130-491-13

Query Match 36.8%; Score 1900; DB 4; Length 608;
Best Local Similarity 49.3%; Pred. No. 6e-150;
Matches 339; Conservative 114; Mismatches 146; Indels 88; Gaps 8;

```
QY 269 VLLLRDRDGPVKVTGNAAALTLRNFCAWKKLNKVSQKHPEYWDTAILFTQDLGATICD 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 ILVHDEQKGPVTSNAALTTLRNFCAWKKLNKVSQKHPEYWDTAILFTQDLGATICD 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 329 TLGMADVMTCDPKRSCSVIEDDGLPSAFTTAHELGHVFNHMDNVKVCVEEFGKLRANH 388
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 TLGMADVMTCDPKRSCSVIEDDGLQAAFTTAHELGHVFNHMDNVKVCVEEFGKLRANH 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 389 MMSPTLLQIDRANPWSACSAAIITDFDLSGHCGLLDQSPKPISLPEDLGASTYLSQ 448
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 MMASMLNLSHDSQSPWSPCSAYMTISFLDNHGECLMDKPNQIQLPDGLPGTSYDANRQC 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 449 ELAFGVSGSKPCP--YMYCTKLWCTGKAGQMVQCTRHFPWADGTCGEGKLCIKGACVER 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 184 QFTFGEDSKHCPDAASTCTLWCTGSGVLVQCTRHFPWADGTCGEGKLCIKGACVER 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 508 HNLNKH---RVDGSNAKWDPYGPCSRTCGGVLQARRQCTNPTPANGKGYCEGVRYKRS 564
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 244 TD-RKHEDTFFHSGWGMWGPWGDSCRTCGGQVQYTRMRECDNPVPKNGKGYCEGVRYKRS 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 565 CNLEPCPSSASGSKSFRREEOCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCLICRANG 624
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 CNLEDCPDN--NGKTFREEOCEAFNEFSKASFGSGPAVEWPKYAGVSPKDRCKLICQAG 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 625 TGYFYVYAPKVVVDGTLCSPDSTSVQVQKCIKAGCDNGLSKKRFDKCGVCGGDNKSCXK 684
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 IGYFFVLQPKVVDGTPCSPDSTSVQVQGVQCVKAGCDRIIDSKKKFKDCGCGGNGSTCKX 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 685 VTGLTKPMHGYVNVVAIPAGASSIDIROGYKGLIGDNNYLALNKSQKYLINGHFVVS 744
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 ISSVTSKAPGYHDIITIPIGATNIEVKORNGSRNNGSFLAKAADGTIYLNGDYILS 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 745 AVERDLVWKSLLRYSGTGTAVESLQASRPILPTEVLSVSGKMTPPRVRYSFYLPKEP 804
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 482 TLEQDWMYKGVVLYRSGSSAALERISFSPLEPLIQVLTVGNALRPKIKTYFYVKKX- 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 805 REDKSHHPKDPGCPVSLHNSVLSLQVQEPDPPRPARVAGSWGSPCSASCGSLQKFAV 864
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 ---KES-----FNAITFS----- 551
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 865 DCRGSAGQRTVPACDAAHRPVETQACGEPCTWELSAWSPCSKSGCGFQRLSKCVGJG 924
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 552 -----AWVEEGECSKTCGKGYKRSKLSCLSD 580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 925 GRLLARDONLHRKPOE--LDFCVLRPC 950
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 581 GGVLSHESCDPLKKPKHFDICTMABC 607
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 5

US-09-369-364A-2

; Sequence 2, Application US/09369364A

; Patent No. 6391610

; GENERAL INFORMATION:

; APPLICANT: Apte, Suneel

; APPLICANT: Hurskainen, Tiina L.

; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

; FILE REFERENCE: 26473/4007/10-30-00

; CURRENT APPLICATION NUMBER: US/09/369,364A

; CURRENT FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 930

; TYPE: PRT

; ORGANISM: mus musculus ADAMTS-5

; US-09-369-364A-2

Query Match 35.9%; Score 1854; DB 4; Length 930;

Best Local Similarity 43.1%; Pred. No. 7.6e-146;

Matches 369; Conservative 116; Mismatches 265; Indels 106; Gaps 13;

Qy 90 GLTGGSDLRRCFYSGVGNVNEPDSFAAVSLCGLRGAFGRGAAYVLSLPLNASAPAAQR 149

Db 121 GLSASSGRHGHCFYRGTVGDSRSLAVFLDGLGDFFAVKHARYTLKPLRGSWAEYER 180

Qy 150 -----NSQGAHLLORRG-----VPGSGDPTSRGVSAGWNPAILRALDPY 191

Db 181 IYDGSRSKILHYNREGSFALPPRASCETPASP----- 217

Qy 192 KPRRAGFGESESRSSG-----RAKRFVSIERYVETILVAD 227

Db 218 -PQESPVSHSRRRSALAPQLLDHSAFSPGNAGPQTWNRRRRSISRARQVELLIVAD 276

Qy 228 ESMVKFHGADLEHLLILLATARLYRHPSTILNPINIVVYKVVLLLRDRDSGPKVTGNAAL 287

Db 277 SSMARMYGRGLQHLVLTMASTANRLSHASTENHRLAVKVVVLLIDKDTLSLEVSKNAAT 336

Qy 288 TLRNFCWOKLKNVSKDKPYWDATILFTRODLGATTCDTLGMADVGTMCDPKRSQSV 347

Db 337 TLNFKCWQHQNQLGDDHEHYDAAILFTREDLCGHSHCDTLGMADVGTICSPERSCAV 396

Qy 348 IEDDGLPSAFTAHLEHGVFNHNDNVKVEEVEFGKLRANHMSPPTLIQIDRANPWSACS 407

Db 397 IEDDGLHAAFTVAHEIGLLGLSHDDSKFCBENFGTTEDKRLMSSILTSIDASKPWSKCT 456

Qy 408 AAILTDFLDGSHGCDLLDQPSKPSLSDPLPGASYTLSSQCELAFGVSGKICPYMQVCTK 467

Db 457 SATITEFLDDGHGNCLLDLPKQILGLPEELPGQYIDATQCCNLTFGPPEYVCPGMDVCAR 516

Qy 468 LWCCTGKARKQMVCTRHFPWADGTSCGEGKLCGLKACVER-----HNLNKHVRVDGSWAKW 522

Db 517 LWCAVVRGQMVCTLTKLPVAVGTCGKGRVCLQCKVDKTKKYYSTSSH---GNNGSW 573

Qy 523 DYPGCSCTCGGQVLARRQCTNTPANGKYCEGVKRVKRSNLEPCPSASCKSREE 582

Db 574 GPWGQSRSCGGVQFAYRHCNAPRNSGRYCTGKRAIYRNSCVTPCP--PNGKSPRHE 631

Qy 583 QCEAFGNHSTNRLTLAVAMPKYSVSPDKCKLICRANGTGYFVVLAPKVVVDGTLC 642

Db 632 QCEAKNGVQSDAKGVKTVFVWPYAGVLPADVCKLICRAGTGYVVPKVTIDGTECR 691

Qy 643 PDSTSVCCVQGCICAKGDCNGLSKRRFDKCGVCGGDNKSKKVTGLFTKPMHGVNFVAI 702

Db 692 PYSNSVCRGRVTRGCDGIIGSLQYDKCGVCGGDNSSCTKIIGTFNKRSKGYTDVRI 751

Qy 703 PAGASSIDIRQYKGLIGDDNYLALKNSQKYLINGHFVVSAYERDLVKVGSLLRYSGT 762

Db 752 PEGATHIKVROPKADQTRFPAYLALKKKTGEYLINGKYMISTSETIIDINGTVWNTSGW 811

Qy 763 GRAVESL-----QASRPITLPTVEVLSVGRKMTPPRVRYSFYLPKPRELKS-----SHPKD 814

Db 812 SHRDQFLHGMGYSATKEI---LIVQILATDPTKALGVRYSFVFPKKTQKVNSVISHGSN 868

Qy 815 PRGPSVLHNSVLSLSNQVEQDDPPRPPARWVAGSWGPGSCSGGLQKRAVDCRGSAGQRT 874

Db 869 KVGPP---HSTQL-----QWVTGPMACSRCTCDTGHHTRTVQCQ-DGNRKL 909

Qy 875 VPACDAAHRPVETQAC 890

Db 910 AKGCLLSQRPSAFKQC 925

RESULT 6

US-09-369-364A-13

; Sequence 13, Application US/09369364A

; Patent No. 6391610

; GENERAL INFORMATION:

; APPLICANT: Apte, Suneel

; APPLICANT: Hurskainen, Tiina L.

; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

; FILE REFERENCE: 26473/4007/10-30-00

; CURRENT APPLICATION NUMBER: US/09/369,364A

; CURRENT FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 1882

; TYPE: PRT

; ORGANISM: Homo sapiens ADAMTS-9

; FEATURE:

; NAME/KEY: MOD_RES

; LOCATION: (468)

; OTHER INFORMATION: Xaa = C

; NAME/KEY: MOD_RES

; LOCATION: (521)

; OTHER INFORMATION: Xaa = Y

US-09-369-364A-13

Query Match

Best Local Similarity 35.8%; Score 1848.5; DB 4; Length 1882;

Best Local Similarity 39.6%; Pred. No. 6.1e-145;

Matches 396; Conservative 144; Mismatches 340; Indels 121; Gaps 23;

Qy 22 EREVVVPIRLDP-----DINGRRYYW-----RGPEDSGDQGLIFQITAFQ 61

Db 49 EYEIVSPIKVNALGEPFFPNVHFKTRRSINSATDPWAFASSSSSTSSQAHYRLSAFG 108

Qy 62 EDFYHLTLTPDAQFLAPAFSTEHLGVP-----LQGLTGGSSDLRRCFYSGVGNVNEPDSRAV 117

Db 109 QQFLENLTACAGFIAPLFTVTLTGTPGVNQTIFYSEEAELKHCFYK----- 155

Qy 118 SLCCGLRGAFGRGAIFYVLSPLPNASAPAAQNSQGAHLLQRRGVPGSGDPTSRGVA 177

Db 156 RLCQYQLRAHG-----RHQPLLREHNRHSKDKKTRAKWGERINLAGDV 202

Qy 178 SGWNPAL-LRALDPYKPRRAGFGESESRSSRRSRRAKRFVSIPIRYVETLVVADSWYKFIHA 236

Db 203 AALNSGLATEARSAYGNKTDNTRKTRHRT---KRFLSYPRFEVLVVDNRMYSYHGE 259

Qy 237 DLEHYLLTLTAARLYRHPSTILNPINIVVYKVVLLLRDRDSGPKVTGNAALTLRNFCAWQ 296

Db 260 NLQHYLTILMSIVASIKDPSIGNLINIVNLVHNEQDGPFSIFNAQTTLKNFCOMQ 319

Qy 297 KLKLVKSDKHPEWDTAILFTRODLGCA--TTCDTLGMADVGTMCDPKRSQSVIEDDGLPS 355

Db 320 HSNPSGGIHH-----DTAVLLTRDIDICRAHDKCDTGLGAEGLTICDPRSCSISEDSGLST 375

Qy 356 AFTTAHELGHVFNHNDNVKVEEVEFGKLRANHMSPPTLIQIDRANPWSACSAAITDPL 415

Db 376 AFTIAHELGHVFNHNDNVKVEEVEFGKLRANHMSPPTLIQIDRANPWSACSAAITDPL 434

Qy 416 DSHGBCDCLLDOP--SKFISLIPEDLPGASYTLSSQCELAFGVSGKPYMQVCTKLWC---T 471

Db 435 DTGYGCELNEPESRYPPLVQPLVGLILYNNKQXELIFGFGSOVCPYMMQCRRLMNCNNV 494

QY 472 GKAKGMVQTRHPFADGTSCEGKLCACACVERHNKLRVDGSAWAKWDPYGPCSRT 531
Db 495 GVHKG--CRQHTPWADGTECEPGKCKXGFCVPK-EMDVPYTDGWSGSPFPGTCSRT 550
QY 532 CGGQVQLARRCTNPTPANGKCYCEGVYKRSNLEPCSPSSASGSKFEEQCEAFNGYN 591
Db 551 CGGQVQLARRCTNPTPANGKCYCEGVYKRSNLEPCSPSSASGSKFEEQCEAFNGYN 591
QY 592 HSTNRLTLAVAWPKYSGVSPRDKCLICRANGTGYFVLAPKVVDTLCSPDSTSCVQ 651
Db 609 FNINGLLENVWPKYSGILMKDKCLFCRVAGNTATYQURDVIDTGPCQDNDICVQ 668
QY 652 GKCTKAGDGNLCKKRFKDCGCGGDKNCKKVTGLFTXPMHGYNFVVAIPAGASSIDI 711
Db 669 GLCQAGCDHVLNKKARRDKCGVCGDSSCKTVAGTFNVHYNTVVAIPAGATNIDV 728
QY 712 RQGYKGLIGDNDYALNKGKYLNGHFWVSAVERDLVYKSLRLYSCTGTAVESLQA 771
Db 729 ROHFSFGTDDNDYALSSKGEFLNGFNVTMAKREIRIGNAVVEYSGETAVERINS 788
QY 772 SPILLEPLTVVLSVGMKTPPVRYSEFLPKPEPREDK---SSHFKDPRGP----- 818
Db 789 TDRIQELLOVLSVGLKLYNEDVRYSEFLPKPEPREDK---SSHFKDPRGP----- 818
QY 819 -----SVLHNSVLSNQVEQDDRRP-----ARWVAGSWGSCSACGS 857
Db 844 ERKRLVCTRESQDLTVS---DQCDRLPOPHITEPCGTCGCDLRLWHVASKSCSACQGL 900
QY 858 GLOKRAVDCR-----GSAGORTVPACDAARHPVETQACGECPT--WLSAWSKSCS 909
Db 901 GYRTLDIYCAKYSRLDGKTEKVDGFCSSHPKPSNRKCSGECNTGWRISAWTECSKSC 960
QY 910 GRGFORSLKCVGHGRLRLARDNLHRRKQELDFCVLRPC 950
Db 961 DGGTORRAICVNTNRDVLDDSKCT-HQEKVTIQRCEEAC 1000

RESULT 7

US-09-122-126B-15
; Sequence 15, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122.126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-122-126B-15

Query Match 35.8%; Score 1847.5; DB 4; Length 930;
Best Local Similarity 44.0%; Pred. No. 2.7e-145;
Matches 392; Conservative 113; Mismatches 291; Indels 95; Gaps 22;
QY 48 SGOGLFIQITAFQEDFYHLITPDQAFSLAPFSTEHLGVPLQGLTGSSJLR---RCFY 104
Db 82 SGGKVGVLVYAGGRFLDLERDGVGVIAGF-----VAGG--GTSAPWRHRSCHF 133
QY 105 GDNVNEPDSFRAVSLCGGLRGAFGYRGAEEYISPL---PNASAPAAORNSQGA---HLL 157
Db 134 GTVDASPRSLAVFDLCGLDGLDFFAVKHARYTLKPLLRGPMAEERKGRVYDGSARTLHV 193
QY 158 QRRG-----VPGGPG-----GDPTSRCKGVASGWNPAIIL--RALDPYKP 193
Db 194 TREGSFALPRASCETPASTPEAHEHAPAHNSPGRALAS-----QLDQSALSP--- 246
QY 194 RRAGFES---RSRRSRGRKRFVSIPIRYVETLVVADESVMVFKHGLHLYLTLLATAA 250

Db 247 -AGSGPQTWRRRRRS-----ISRARQVELLVADASMARLYGRQLHYLLTASIAN 299
QY 251 RLYRHPSTILNPINIVVYKVLRLDRSGPKVTGNAATLLENFCAMQKLNKYSGKHPEY 310
Db 300 RLYSHASTENHIRUAVVYKVLRLDRSGPKVTGNAATLLENFCAMQKLNKYSGKHPEY 310
QY 311 DTAILFTRQDLCCGATTCTDLMADVGTMCDPKRSKSVIEDDGLPSAFTTAHELGHVFNMP 370
Db 360 DAAILFTRQDLCCGATTCTDLMADVGTMCDPKRSKSVIEDDGLPSAFTTAHELGHVFNMP 370
QY 371 HDNVKVEEYFGKLRANHMSPILIQIDRANPWSASAAIITFDLSGHDGCLLDQSPXP 430
Db 420 HDSKFCBETFGSTEDRKLMSILTSIDASKPMKCTSATITEFLDDGHGNCILDLPRXO 479
QY 431 ISLPEDLPASYYTLLSQOCELAFGVGSKPCPYMOYCTKLWCTGKAKGMVQTRHPFADG 490
Db 480 ILGPEELPGQTYDATQCNLTFFGEYSVCGMDVCARLACAVVYRQGMVCLTKKLPAVE 539
QY 491 TSCGEGKLCACACVER-----HNLNKHVRDVGSAWAKWDPYGPCSRTCGGQVQLARR 545
Db 540 TPCGKGRICLQCKVDKTKKKYSTSSH--GNMWSGWSGQCSRSQGGVQVAFYRHCNN 596
QY 546 PTBANGKCYCEGVYKRSNLEPCSPSSASGSKFEEQCEAFNGYHSTNRLTLAVAWVP 605
Db 597 PAPRNNGRICYTKRAIRNSCSLMPCP--PAGKSRHEQCEAKNGYOSDAKGVYTFVWVP 654
QY 606 KYSVSPRDKCLICRANGTGYFVLAPKVVDTLCSPDSTSVYQVQKCIKACGDNGLS 665
Db 655 KYAGVLPADVCKLACRAKGTGYVVFSPKVTGTECPYNSVSVCRGKCVRTCDGLIGS 714
QY 666 KKRDEKCCVCGGDKNCKKVTGLTKPMHGYNFVVAIPAGASSIDIQRQYKGLIGDNY 725
Db 715 KLOYDKCVCGGDSSCTKIVGTENKSKGYTDVVRPEGATHIKVRQFAKQDTFTAY 774
QY 726 LALKNSOCKYLLNGHFVYSAVERDLVYKSLRLYSCTGTAVESL-----QASRFLLEPL 780
Db 775 LALKNGEYLLNGHYMISTETIIDINGVYVNYSGHSHRDEFLGNGYSATKEL---LI 831
QY 781 VEVLVSGMKTTPPVRYSEFLPKPEPREDKSHHPKDPGSPVLHNSVLSL-SNQVEQDDRP 839
Db 832 VQILATDPTKLDVRYSEFLPKPEPREDKSHHPKDPGSPVLHNSVLSL-SNQVEQDDRP 839
QY 840 PARWVAGSWGSCSACSGSLQKRAVDCRGSGAGORTVPACDAARHPVETQAC 890
Db 878 --QWVTPWLAACRSTCDTGHTRTVQCO-DGNRKLAKGCLPSQRPSAFKQC 925
RESULT 8
US-09-369-364A-15
; Sequence 15, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369.364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-9
US-09-369-364A-15

Query Match 34.8%; Score 1798.5; DB 4; Length 874;
Best Local Similarity 42.1%; Pred. No. 3e-141;
Matches 369; Conservative 122; Mismatches 302; Indels 83; Gaps 18;
QY 115 AAVSLCGGLRGAFGYRGAEEYISPLPNASAPAAORNSQGAHLRQVPGGSGDPTSR 174


```
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-8

Query Match
Best Local Similarity 24.7%; Score 1273; DB 4; Length 481;
Matches 236; Conservative 87; Mismatches 141; Indels 38; Gaps 12;

QY 402 PWSACSAIITDFLDSHGDCCLLDQPSKPISLPDLPCASYTLSSQOCELAFAFGVSKGPCP- 460
Db 7 PWSPCSAYVLTBLDDGHGDCLLD-----DGHSTLYELDDQCKQIFGPDHRCPN 56

QY 461 --YMOYCYKLNCTGKRAKQOMVCOIRH--FPWADGTSCEGKCLKAGACVFNHNLN--KHR 514
Db 57 TSVEDICVOLWCRHRDSDEPCHTKNALLWADGTPCGHLCUDGSCVIREVENPKAV 116

QY 515 VDSWAKWDPYPCSTRCTCGGQVQLARRQCTNPTPANGKCYCEGVYRVYRSCNLEPCPSSA 574
Db 117 VDDMGWPGWPGWQCSRTCTCGGQVQFESNRECDNPAPQNGRFLGERVKYQSKTECP--P 174

QY 575 SKSREDOCEAFNGYNHSTNRLTLAVAVPKYSGVSPDRCKKLCIRANGTGYFYVLAPK 634
Db 175 NGKSFREDOCEKYNAYNH--TLDGNGFLQWVPKYSGVSPDRCKLFCRAR--KSEFKVFETK 233

QY 635 VVDGTLCSPDSTSVQVQKCKTACAGDGNLGGKRRFKDGCVCVGGDNKCKAVTGLTKPMH 694
Db 234 VIDGTLGPDYLAICVRQCYKAGCDHVNSPKKLDKCGVCGGKTACRVSGSFTFESY 293

QY 695 GTFVVAIPAGASSIDIRQYKGLIGDNDNYLALNKNSQGYLLNGHFWGSAVERDLVVK 754
Db 294 GYNDIYVTPAGATNIDVQRSPGQVNDGYSYALKTANGQVLLNGNLAIENAEQDILMK 353

QY 755 SILRYSGTCTAVESLQASRPILPTVEVLSV-GKMTPPRVRYSYFLPKPEPREDKSHPK 813
Db 354 TILKYSGSWATLERLOSQALPELITVQLLVSEVFPKVKYTFVFNPE--TDFNVQSS 411

QY 814 DPGFVSLHNSVLSLSNVEQPDPRPARVAGSWGPCASCGSLQKRAVDCRGSGAGQR 873
Db 412 KERAST--NIQSL-----PYAEWVLGDMSECPSTCGGGWRRR(VECRDPDSQA 458

QY 874 TVPACDAHRPVEYTOAGC-EPC 894
Db 459 S-DTCDALKPEDAKPCGSQPC 479

RESULT 11
US-09-369-364A-22
; Sequence 22, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 22
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-5
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (39)
; OTHER INFORMATION: Xaa = M
US-09-369-364A-22

Query Match
Best Local Similarity 24.5%; Score 1264; DB 4; Length 518;
Matches 244; Conservative 73; Mismatches 185; Indels 36; Gaps 10;

; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-8

Query Match
Best Local Similarity 24.7%; Score 1273; DB 4; Length 481;
Matches 236; Conservative 87; Mismatches 141; Indels 38; Gaps 12;

QY 402 PWSACSAIITDFLDSHGDCCLLDQPSKPISLPDLPCASYTLSSQOCELAFAFGVSKGPCP- 460
Db 7 PWSPCSAYVLTBLDDGHGDCLLD-----DGHSTLYELDDQCKQIFGPDHRCPN 56

QY 461 --YMOYCYKLNCTGKRAKQOMVCOIRH--FPWADGTSCEGKCLKAGACVFNHNLN--KHR 514
Db 57 TSVEDICVOLWCRHRDSDEPCHTKNALLWADGTPCGHLCUDGSCVIREVENPKAV 116

QY 515 VDSWAKWDPYPCSTRCTCGGQVQLARRQCTNPTPANGKCYCEGVYRVYRSCNLEPCPSSA 574
Db 117 VDDMGWPGWPGWQCSRTCTCGGQVQFESNRECDNPAPQNGRFLGERVKYQSKTECP--P 174

QY 575 SKSREDOCEAFNGYNHSTNRLTLAVAVPKYSGVSPDRCKKLCIRANGTGYFYVLAPK 634
Db 175 NGKSFREDOCEKYNAYNH--TLDGNGFLQWVPKYSGVSPDRCKLFCRAR--KSEFKVFETK 233

QY 635 VVDGTLCSPDSTSVQVQKCKTACAGDGNLGGKRRFKDGCVCVGGDNKCKAVTGLTKPMH 694
Db 234 VIDGTLGPDYLAICVRQCYKAGCDHVNSPKKLDKCGVCGGKTACRVSGSFTFESY 293

QY 695 GTFVVAIPAGASSIDIRQYKGLIGDNDNYLALNKNSQGYLLNGHFWGSAVERDLVVK 754
Db 294 GYNDIYVTPAGATNIDVQRSPGQVNDGYSYALKTANGQVLLNGNLAIENAEQDILMK 353

QY 755 SILRYSGTCTAVESLQASRPILPTVEVLSV-GKMTPPRVRYSYFLPKPEPREDKSHPK 813
Db 354 TILKYSGSWATLERLOSQALPELITVQLLVSEVFPKVKYTFVFNPE--TDFNVQSS 411

QY 814 DPGFVSLHNSVLSLSNVEQPDPRPARVAGSWGPCASCGSLQKRAVDCRGSGAGQR 873
Db 412 KERAST--NIQSL-----PYAEWVLGDMSECPSTCGGGWRRR(VECRDPDSQA 458

QY 874 TVPACDAHRPVEYTOAGC-EPC 894
Db 459 S-DTCDALKPEDAKPCGSQPC 479

RESULT 12
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match
Best Local Similarity 32.9%; Score 1248.5; DB 4; Length 1224;
Matches 332; Conservative 126; Mismatches 386; Indels 165; Gaps 38;

QY 51 QGLIFQITAFQEDFYHLHTPPAQFLAPAFSTEHGLGVPLQGLTGSS-----DLRRCFY 103
Db 95 ESHLRLKGRIDFHVLDLRTSSSLVAPGFIQTL-----GKTGKSVQTLPPDEF--CFY 147

QY 104 SGDVAEPDPSFAVSLCGLRGAFYRGAEVYISPLPNASAPAAQNSQG---AHLLQRR 160
Db 148 QGSLSHRNSVALSTCQGLSGMIRTEADYFLRPLPSHLKSLGKRAQGSFVHLYKR 207

QY 161 G----VPGGSGDPYSR-----CGVASGNWPA-----ILRA 187
Db 208 STEPIAPGASEVLYTSRTELAAHQLHSDLLRLGLPQKQHFGRKKYKMPQPPREDLFL 267
```


US-09-369-364A-17

Search completed: May 9, 2003, 15:24:43
Job time : 36 secs

Query Match 22.3%; Score 1152; DB 4; Length 1081;
Best Local Similarity 31.5%; Pred. No. 3.1e-87;
Matches 318; Conservative 120; Mismatches 400; Indels 172; Gaps 35;

QY	64	FYLHLPDAQFLAPAFSTEHGLGVPLQGLTGGSDLRRCFYSGDVNAE-PCGFAAVSLCGG	122
DB	65	FLNLTRSRRLLAGRVSVYEW--TREGAWQRAARPHCLYAGHLOGQAS\$SHVAISTCGG	122
QY	123	LRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGPGSDPTSRGV--ASGW	180
DB	123	LHGLIIVADEEYVLEPLHGGPKGSRSPESGPHCVYKRSLRHPHLD--IACGVREKFPW	180
QY	181	--NPAILRALDPYKPRAGFGESRRRRSRRAKRFVSPRYVETLVVADE\$MVKEHG-AD	237
DB	181	KGRPWMLRTLKP--PPARPLGNETERGQPG-LKRSVSRERYVETMDVADKMWVAYHGRD	237
QY	238	LEHYLLTLATAARLYRHP\$ILNPIVIVVVKVLLLRDRSDGPKVTGNAALILRNECAMOK	297
DB	238	VEQVLAIMNIVAKLFQDSSLGSTVNLVTRLILTEDOPTLEITHHAGKSLDSFCKWCK	297
QY	298	KL-----NKVSDKHPEYWDTAILTTRDLC--GATTCTGLMADVGTMDPKRSCSVI	348
DB	298	SIVNHSHGNAIPENGVAHNTAVLITRYDICIYKKNKPGCTILGLAR-WAECV\$AREAAAS	356
QY	349	EDGLPSAFTTAHELGHVFNPHDNYKVEEYFG-----KLRAHHM\$PTLIQIDRANP--	402
DB	357	MRTLAATSVHHCHEIGHTFGMNHHDGVNCGGARGQDPAKLMAHITMKT-----NP	409
QY	403	WSACSAIITFDLDSHGDCLLDPSK-PISLPEDLP\$GASYTL\$OQCELAFUGVSKPCPY	461
DB	410	WSSCNRDYIT\$FLDSGLGLNNRPQDPVYPTVAPGOAYDADEQCRFHGVKSRQCKY	469
QY	462	MOYCTKLWCTGKAKQMVCTQTRHFPWADGTSCGEGL-----CLKGACVERHNLNKRVDG	517
DB	470	GEVCSLWCLSKNR--CITNSIPAAEGTLQOHTHIDKGWYKRVCP-FC\$RPEGVDG	525
QY	518	SWAKWDPYGC\$RTCGGVOLARRQCTNPTPANGKYKCEGVVKYK\$CNLEPCPSSASGK	577
DB	526	AWGPWTWGC\$RTCGGVSSSRHCDSPRTIGGKYLGERRRH\$RSCNTDDCPPGS--Q	583
QY	578	SFREEQCEAFNG-----YNHSTNRLTLAVAVWPKYSGVSPRDKCKLICKRANGTGYFYV	630
DB	584	DFREVOCAEFD\$IPFRGKYFKWKT\$YR-----GGGVK--ACSLT\$LAEGNFYTE	630
QY	631	LAPKVDGTLCPD\$T\$VCCVQGCIRAGCDNGLSKKRPDKCGVCGGDNK\$CKKVTGLFT	690
DB	631	RAAAVVDGTPCRPDTVIC\$GEC\$KHVGC\$DRVLG\$DLREDKRVCGGDSACETIEGVFS	690
QY	691	--KPMHGYNFVAIPAGASSIDIRQGYKGLIGDNYLALKN\$QKYL\$LLNGHFV\$SAVER	748
DB	691	PASPGAGYEDVW\$IPK\$GVHFIQD-----LNL\$SLH\$ALKGDOESLLLEGLPGTQP	745
QY	749	DLVVK\$GLLRYSGTGTAVESLQASRPIL\$PLTVEVL\$VSGKMT\$PRVRY\$FLP-----	801
DB	746	-LPLAGTTFLQROGPDQV\$LEALGPIN\$LVNVLARTEL--PALRY\$ENAPIARD\$SLP	802
QY	802	-----KEPREDK--SSHPKD	814
DB	803	PY\$WHYAPWTK\$CAQCA\$G\$QVQAVECRNQLD\$SSAVAPHYCSAH\$KLPK\$QACNTEPCP	862
QY	815	P-----RGPSV\$LHNSV\$L\$SLNQVEQ-----PDDRPPA-----	841
DB	863	PDWVVGWNSLCSR\$CDAGVR\$ST\$VVCORRV\$AAE\$KALD\$DSAC\$PQPRPV\$LEACHGPTCP	922
QY	842	-RWVAG\$WGPC\$SAC\$G\$CLOKRAVDC\$R\$SAGORTVPA--CDA\$HRPV\$ETOACG-EP	895
DB	923	PEWATLDW\$ECTP\$CGGLRH\$RVVLC\$K\$SADOR\$TL\$PPCHCL\$P\$AAK\$P\$STW\$RCNLR	982
QY	896	TWEL\$AWS\$PC\$SK\$GGR\$FQR\$R\$K\$CV\$GG\$RLLAR\$DQC\$NLRK\$PQELDFC	945
DB	983	RWVT\$ENGE\$C\$TQ\$CGL\$GQ\$Q\$Q\$RTV\$CT\$HT\$GQ--P\$RECTEALR\$ESTMQOC	1030

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:17:26 ; Search time 44 Seconds
(without alignments)
2877.002 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: I MLLGIITLPAFAGTAGGSE.....DQNLHRKFOHLDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5162	100.0	950	23	AAE22541 Human protease #2.
2	5156	99.9	950	22	AAG62299 Human metalloprotease
3	5117	99.1	952	23	AAU74751 Human protease PRT
4	4862.5	94.2	928	23	AAU72899 Human metalloprotease
5	2497.5	48.4	505	21	AB21257 Rat metalloprotease
6	2486.5	48.2	950	20	RAY49501 Human METH1 protei
7	2486.5	48.2	950	22	AAU73549 Human ADAM-type me
8	2486.5	48.2	950	22	AAB50002 Human METH1. Homo
9	2486.5	48.2	967	19	AAW80285 Human integrin lig
10	2486.5	48.2	968	22	AAB50011 Protein; SEQ ID 12

11	2485.5	48.1	967	20	AAU04142 Human Tango-71 pro
12	2464.5	47.7	967	20	AAU78189 Human secreted pro
13	2448.5	47.4	950	21	AAV53899 Amino acid sequenc
14	2325	45.0	896	21	AB21265 Mouse metalloprote
15	2274	44.1	727	20	AAU78435 Human ADAMTS-1 pro
16	2138.5	41.4	890	20	AAU49502 Human METH2 protei
17	2138.5	41.4	890	22	AAB50003 Human METH2. Homo
18	2129.5	41.3	889	22	AAU74946 Human ADAM type me
19	2115	41.0	905	22	AAU72884 Human ADAMTS-8 am
20	1966	38.1	381	21	AB21261 Human metalloprote
21	1965	38.1	1629	23	ABG30703 Human aggrecanase
22	1965	38.1	1629	23	AAO14448 Human ADAMTS-S1 pr
23	1965	38.1	1916	23	AAE19173 Human protease, PR
24	1965	38.1	1935	23	AAU72896 Human metalloprote
25	1957	37.9	1602	23	ABG30702 Human aggrecanase
26	1943.5	37.7	1073	21	AB21264 Human metalloprote
27	1939.5	37.6	1934	22	AAE72301 Human ADAMTS-9 alt
28	1919	37.2	367	23	AAE22542 Human protease #3.
29	1915	37.1	837	20	AAU75425 Human aggrecan deg
30	1914	37.1	837	21	AAV99429 Human PRO1563 (UNQ
31	1914	37.1	837	22	AAU29199 Human PRO polypept
32	1914	37.1	837	22	AAU29199 Human aggrecanase-
33	1913	37.1	837	22	AAG78228 Human metalloprote
34	1913	37.1	840	21	AB21256 Human metalloprote
35	1903	36.9	680	21	AB21251 Human metalloprote
36	1882	36.5	1907	23	AAU77133 Human protease #12
37	1874	36.3	1505	23	AAU72897 Human metalloprote
38	1872	36.3	947	22	AAB56950 Human metalloprote
39	1854	35.9	930	22	AAU72280 Murine ADAMTS-5 am
40	1848.5	35.8	1882	22	AAU72286 Human ADAMTS-9 am
41	1847.5	35.8	930	20	AAU75426 Human aggrecan deg
42	1840	35.6	929	21	AAU75426 Human OREF ORF90
43	1798.5	34.8	874	22	AAU72287 Murine ADAMTS-9 am
44	1783	34.5	958	21	AB21255 Human metalloprote
45	1762	34.1	870	21	AB21252 Rat metalloprotein

ALIGNMENTS

RESULT 1
AAE22541
ID AAE22541 standard; Protein; 950 AA.

AC AAE22541;

XX 26-JUL-2002 (first entry)

DT Human protease #2.

DE Human; novel human protein; NHP; protease; biological disorder; obesity;

XX high blood pressure; arthritis; connective tissue disorder; infertility;

XX gene therapy; enzyme.

XX Homo sapiens.

XX WO200226949-A2.

XX 04-APR-2002.

XX 27-SEP-2001; 2001WO-US30350.

XX 29-SEP-2000; 2000US-236689P.

XX (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Hilbun E;

XX WPI: 2002-372123/40.

XX N-PSDB; AAB35569.

XX Novel nucleic acid encoding a human protease, useful as a hybridization

XX probe for screening libraries and assessing gene expression patterns -

XX PS Claim 6: Page 36-38; 41pp; English.

CC The present sequence is novel human protein (NHP), human protease.

CC NHPs share structural similarity with animal proteases particularly

CC zinc metalloproteases. Sequences of the invention are useful in

CC therapeutic, diagnostic and pharmacogenomic applications. NHP

CC polynucleotides are used as hybridisation probes for screening

CC libraries and assessing gene expression patterns. They can also be

CC used for treating related biological disorders such as obesity, high

CC blood pressure, arthritis, connective tissue disorders and infertility.

CC They are also used in gene therapy.

XX XX Sequence 950 AA;

SQ Query Match 100.0%; Score 5162; DB 23; Length 950;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

DB 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

QY 61 QEDFYHLTPDQAFAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDSFAAVSLC 120

DB 61 QEDFYHLTPDQAFAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDSFAAVSLC 120

QY 121 GGLRGAFYGAIEYVISPUNASAPAAQNSOGAHLQRRGVGGPDSRCGVASGW 180

DB 121 GGLRGAFYGAIEYVISPUNASAPAAQNSOGAHLQRRGVGGPDSRCGVASGW 180

QY 181 NPAILRALDPKPRRAGFSGRSRRSRGAKRFVSIPRYETLVVADESVMVRFHGDLEH 240

DB 181 NPAILRALDPKPRRAGFSGRSRRSRGAKRFVSIPRYETLVVADESVMVRFHGDLEH 240

QY 241 YLTLTATAARLYRHPISILNPIVIVVVKLLDRDSGPKVTGNAALTLENFCAMOKKLN 300

DB 241 YLTLTATAARLYRHPISILNPIVIVVVKLLDRDSGPKVTGNAALTLENFCAMOKKLN 300

QY 301 KVSQKHEPYWDITALLTRQDLCCATTCDTLGMADVGTMDCKPRSCSVIEDDGLPSAFTTA 360

DB 301 KVSQKHEPYWDITALLTRQDLCCATTCDTLGMADVGTMDCKPRSCSVIEDDGLPSAFTTA 360

QY 361 HELGHVFNMPHDNVKVEEYFVGLKLANHMSPTLIQIDRANPWSACSAALITDFLDSHG 420

DB 361 HELGHVFNMPHDNVKVEEYFVGLKLANHMSPTLIQIDRANPWSACSAALITDFLDSHG 420

QY 421 DCLLDQSKPISLPEDLPGASYTLQOCELAFGVSKPCPYMOYCTKLWCTGKAKQWVC 480

DB 421 DCLLDQSKPISLPEDLPGASYTLQOCELAFGVSKPCPYMOYCTKLWCTGKAKQWVC 480

QY 481 QTRHFPWADGTSCGEGKLCILKGCVERHNLKRVDSWAKWDPYPCSFICGGGVQLAR 540

DB 481 QTRHFPWADGTSCGEGKLCILKGCVERHNLKRVDSWAKWDPYPCSFICGGGVQLAR 540

QY 541 RQCTNPTPANGKYCEGVVRVYKSCNLEPCSSASGKSPREOCEAFNGYHNHSTNLTLA 600

DB 541 RQCTNPTPANGKYCEGVVRVYKSCNLEPCSSASGKSPREOCEAFNGYHNHSTNLTLA 600

QY 601 VAWPKYSGVSPDKCKLICRANGTGYYVLAQVVDGTLCSPDSTSVCGQKICRAGCD 660

DB 601 VAWPKYSGVSPDKCKLICRANGTGYYVLAQVVDGTLCSPDSTSVCGQKICRAGCD 660

QY 661 GNLGSKRFRDKGCVGGDNKCKKVTGLFTKPMHGYNFVAIPAGASSIDIRQGYKGLI 720

DB 661 GNLGSKRFRDKGCVGGDNKCKKVTGLFTKPMHGYNFVAIPAGASSIDIRQGYKGLI 720

QY 721 GDDNYLALKNSQGYLLNGHFVSAVERDLVWKSILLYSGTGTAVESLQASRILLEPLT 780

DB 721 GDDNYLALKNSQGYLLNGHFVSAVERDLVWKSILLYSGTGTAVESLQASRILLEPLT 780

QY 781 VEVLVSGKMTPPRVRYSYFLPKPEDEKSSHPKDPGRGSPVHLNSVLSLNOVEQPDORPP 840

DB 781 VEVLVSGKMTPPRVRYSYFLPKPEDEKSSHPKDPGRGSPVHLNSVLSLNOVEQPDORPP 840

QY 841 ARWVAGSWGSCSGSLQKRAVDCRGSGAGQRTVPACDAHRPVTQACGECPTWELS 900

DB 841 ARWVAGSWGSCSGSLQKRAVDCRGSGAGQRTVPACDAHRPVTQACGECPTWELS 900

QY 901 AWSPCSKSCGGRGFRRLSKVCGHGRLLARDQCNLHHRKPOELDFCVLRPC 950

DB 901 AWSPCSKSCGGRGFRRLSKVCGHGRLLARDQCNLHHRKPOELDFCVLRPC 950

RESULT 2

AAG62299

ID AAG62299 standard; protein; 950 AA.

XX AC AAG62299;

XX AC AAG62299;

DT 23-AUG-2001 (first entry)

XX Human metalloprotease MDTs6 protein.

DE Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;

KW osteopathic; antiarthritic.

KW Homo sapiens.

OS WO200134785-A1.

PN 17-MAY-2001.

PD 10-NOV-2000; 2000WO-JP07917.

XX 11-NOV-1999; 99JP-0321740.

PR 16-NOV-2000; 2000JP-0144020.

XX (YAMA) YAMANOCHI PHARM CO LTD.

PA (KAZU-) KAZUSA DNA RES INST.

PI Yamaaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

DR WPI; 2001-343602/36.

DR N-PSDB; AAH41003.

XX Metalloprotease with aggrecanase activity for treating joint diseases especially osteoarthritis

PS Claim 1; Page 56-60; 85pp; Japanese.

XX This invention relates to a metalloprotease with aggrecanase activity.

CC The invention includes protein and DNA sequences of the metalloprotease,

CC vectors containing the DNA, host cells transformed by the vectors, and

CC antibodies directed against the metalloprotease. The antibodies, protein

CC and DNA sequences can be used in the treatment and prevention of joint

CC diseases, particularly osteoarthritis. The treatment may result in

CC osteopathic and antiarthritic activity. The present sequence represents

CC the metalloprotease of the invention termed MDTs6.

XX XX Sequence 950 AA;

SQ Query Match 99.9%; Score 5156; DB 22; Length 950;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

DB 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

QY 61 QEDFYHLTPDQAFAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDSFAAVSLC 120

DB 61 QEDFYHLTPDQAFAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDSFAAVSLC 120

QY 121 GGLRGAFYGAIEYVISPUNASAPAAQNSOGAHLQRRGVGGPDSRCGVASGW 180

Db 121 GGLRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGSPGDETSRCGVASGW 180
QY 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESXKPFHGADEH 240
Db 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESXKPFHGADEH 240
QY 241 YLLTLLATAARLYRHPSTILNPINIVVVKLLRRDRSGPKVTGNAALTLRNFCAWQKLN 300
Db 241 YLLTLLATAARLYRHPSTILNPINIVVVKLLRRDRSGPKVTGNAALTLRNFCAWQKLN 300
QY 301 KVSCKHPYWDAILFTTRQDLGGATCTDGLMADVGTCMDPKRSCSVIEDGLPSAFTTA 360
Db 301 KVSCKHPYWDAILFTTRQDLGGATCTDGLMADVGTCMDPKRSCSVIEDGLPSAFTTA 360
QY 361 HELGHVFNHNDNVKVCBEVEFKLRANHMSPTLIQIDRANPWSACSAALITDFLDSHG 420
Db 361 HELGHVFNHNDNVKVCBEVEFKLRANHMSPTLIQIDRANPWSACSAALITDFLDSHG 420
QY 421 DCLLDQPSKPIISLPEDLPASYYTISQOCELAFGVGSKPCPYMYCTKLWCTGKAKGMVC 480
Db 421 DCLLDQPSKPIISLPEDLPASYYTISQOCELAFGVGSKPCPYMYCTKLWCTGKAKGMVC 480
QY 481 QTRHFPWADGTSCEGKLCILKAGACVERHNLNKHVRDGSWAKWDYPGCSSTCGGQVQLAR 540
Db 481 QTRHFPWADGTSCEGKLCILKAGACVERHNLNKHVRDGSWAKWDYPGCSSTCGGQVQLAR 540
QY 541 ROCTNPTPANGKYCEGVYRKYRSCNLEPCPSSASGKSFREEOCEAFNG:NHSTNRLTLA 600
Db 541 ROCTNPTPANGKYCEGVYRKYRSCNLEPCPSSASGKSFREEOCEAFNG:NHSTNRLTLA 600
QY 601 VAWPKYSGVSPDKCKLICRANGTGYFYVLAPKVDGTLCSDDSTSVGVGCKIKAGCD 660
Db 601 VAWPKYSGVSPDKCKLICRANGTGYFYVLAPKVDGTLCSDDSTSVGVGCKIKAGCD 660
QY 661 GNLGSKKRFKDCGCGGDKNSCKKVTGLFTKPMHGYNFVVAIPAGASSILIRORGKGLI 720
Db 661 GNLGSKKRFKDCGCGGDKNSCKKVTGLFTKPMHGYNFVVAIPAGASSILIRORGKGLI 720
QY 721 GDDNYLALKNSQKYLNGHFVSAVERDLVWKSLLRYSCTGTAVESL:ASRPILLEPT 780
Db 721 GDDNYLALKNSQKYLNGHFVSAVERDLVWKSLLRYSCTGTAVESL:ASRPILLEPT 780
QY 781 VEVLVSKMTPPRVRYSEYLPKPREDKSSHDPKPCPSVLHNSVLNSQVQPDORPP 840
Db 781 VEVLVSKMTPPRVRYSEYLPKPREDKSSHDPKPCPSVLHNSVLNSQVQPDORPP 840
QY 841 ARWVAGSWGPCSCGSLQKRAVDCRGSAQRTVPACDAHRPVTQACGECPCPTWELS 900
Db 841 ARWVAGSWGPCSCGSLQKRAVDCRGSAQRTVPACDAHRPVTQACGECPCPTWELS 900
QY 901 AWSPCSKSGRGFORSLKCVGHGRLRLARDQCNLHRKPQELDFCVLRPC 950
Db 901 AWSPCSKSGRGFORSLKCVGHGRLRLARDQCNLHRKPQELDFCVLRPC 950

RESULT 3

ID AAU74751 standard; Protein; 952 AA.
XX AAU74751;

AC AAU74751;

XX 09-APR-2002 (first entry)

DE Human protease PRTS-11 protein sequence.

XX Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.
XX Homo sapiens.

XX WO200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US19178.

XX 16-JUN-2000; 2000US-212336P.

XX 22-JUN-2000; 2000US-213955P.

XX 29-JUN-2000; 2000US-215396P.

XX 07-JUL-2000; 2000US-216821P.

XX 14-JUL-2000; 2000US-218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;
PI Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalla A, Khan FA;
PI Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;

XX WPI; 2002-090437/12.

XX N-PSDB; ABK12894.

XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
in the diagnosis, treatment and prevention of gastrointestinal (e.g.
gastritis), cardiovascular (e.g. atherosclerosis) and cell
proliferative (e.g. cancer) disorders -

XX Claim 1; Page 144-146; 177pp; English.

XX The present invention relates to twenty one new human proteases,
referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
polypeptides of the invention are useful in the diagnosis, treatment and
prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
myocardial infarction, autoimmune/inflammatory e.g. acquired
immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
proliferative e.g. cancer, developmental e.g. Duchenne and Becker
muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
epilepsy and Alzheimer's disease and reproductive e.g. infertility and
endometriosis disorders. Numerous other examples of each disorder are
given in the specification. The present protein sequence represents
the human protease PRTS-11 protein of the invention.

XX SQ Sequence 952 AA;

Query Match 99.1%; Score 5117; DB 23; Length 952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 947; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 MLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYWRGPDGGLIFQITAF 60

Db 1 MLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYWRGPDGGLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLOGLTGSSDLRRRCFYSGDVNAEPDSEAAVSLC 120

Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLOGLTGSSDLRRRCFYSGDVNAEPDSEAAVSLC 120

QY 121 GGLRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGSPGDTSRGCVASGW 180

Db 121 GGLRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGSPGDTSRGCVASGW 180

QY 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESXKPFHGADEH 240

Db 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESXKPFHGADEH 240

QY 241 YLLTLLATAARLYRHPSTILNPINIVVVKLLRRDRSGPKVTGNAALTLRNFCAWQKLN 300

Db 241 YLLTLLATAARLYRHPSTILNPINIVVVKLLRRDRSGPKVTGNAALTLRNFCAWQKLN 300

QY 301 KVSCKHPYWDAILFTTRQDLGGATCTDGLMADVGTCMDPKRSCSVIEDGLPSAFTTA 360

Db 301 KVSCKHPYWDAILFTTRQDLGGATCTDGLMADVGTCMDPKRSCSVIEDGLPSAFTTA 360

Nucleic Acid
132 7473089CBI
98.8% nu Acid
100% Q.A.
SEA 20 NO 11
7473089CBI

QY 541 RCTNPTPANGKYCEGVVYKRSNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
|||||
Db 566 RCTNPTPANGKYCEGVVYKRSNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 625
QY 601 VAWPKYSGVSPRDKKLCICRANGTGYFVLAPK-VVDGTLCSPDSTSVQVQKCIKAGC 659
|||||
Db 626 VAWPKYSGVSPRDKKLCICRANGTGYFVLAPK-VVDGTLCSPDSTSVQVQKCIKAGC 685
QY 660 DGNLGSKKRFDKCGVCGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRORYKGL 719
|||||
Db 686 DGNLGSKKRFDKCGVCGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRORYKGL 745
QY 720 IGDNDYLAALKNSOGKYLKNGHFWVSAVERDLVVKGSLLRYSGTGTAVESIQASRPILLEPL 779
Db 746 IGDNDYLAALKNSOGKYLKNGHFWVSAVERDLVVKGSLLRYSGTGTAVESIQASRP----- 800
QY 780 TVEVLVSGKWTTPRVRYSYFLKPEFREDKSSHPKDPGSPVLSLSNQVEQPDPRP 839
Db 801 -----NSVLSLSNQVEQPDPRP 817
QY 840 PARWAGSWGSPCSASGSGLOKRAVDRCGSAGORTVPACDAHHRVETQACGECPTWEL 899
Db 818 PARWAGSWGSPCSASGSGLOKRAVDRCGSAGORTVPACDAHHRVETQACGECPTWEL 877
QY 900 SAWSPCKSCGRGFORRSKCVGHGRLRLARDQCNLHRKPQELDFCVLRQC 950
Db 878 SAWSPCKSCGRGFORRSKCVGHGRLRLARDQCNLHRKPQELDFCVLRQC 928

RESULT 5
AAB21257
ID AAB21257 standard; Protein; 505 AA.
XX AC AAB21257;
XX DT 23-FEB-2001 (first entry)
XX DE Rat metalloproteinase ADAMTS-5.
XX KW Rat; ADAMTS-5; metalloproteinase; ADAM;
KW a disintegrin and metalloproteinase domain; thrombospondin domain;
KW vaccine; nontropic; neuroprotective; antiparkinsonian;
KW cerebroprotective; cytostatic; antiarthritic; immunosuppressive;
KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;
KW autoimmune disease; brain tumour; brain injury.
XX OS Rattus norvegicus.
XX PN WC2000053774-A2.
XX PD 14-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US06237.
XX PR 08-MAR-1999; 99US-0264585.
XX PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
XX PI Keiner GS, Clark M, Maki RA;
XX WPI: 2000-594326/56.
XX DR N-PSDB; AAA95827.
XX PT Polynucleotide encoding novel members of a disintegrin,
PT metalloproteinase and thrombospondin domain protein family used to
PT prevent and treat Alzheimer's disease, cancer and autoimmune diseases -
XX Claim 12; Fig 14; 129pp; English.
XX CC The present sequence is rat metalloproteinase ADAMTS-5. The
CC ADAMTS family of proteins is closely related to the ADAM (A Disintegrin
CC and Metalloproteinase Domain) family. Members of the ADAMTS family

CC contain a thrombospondin domain in addition to the disintegrin and
CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are
CC useful for the manufacture of medicaments for treating conditions
CC associated with neuroinflammation and/or neurodegeneration, such as
CC Alzheimer's disease, Parkinson's disease and stroke. They are also
CC useful for treating conditions associated with cell proliferation, cell
CC migration, inflammation and/or angiogenesis, such as cancer, arthritis
CC and autoimmune diseases. They can be used to treat patients afflicted
CC with an invasive tumour, a brain tumour or brain injury.

XX Sequence 505 AA;

Query Match 48.4%; Score 2497.5; DB 21; Length 505;
Best Local Similarity 93.5%; Pred. No. 4.5e-185;
Matches 462; Conservative 8; Mismatches 21; Indels 3; Gaps 2;

QY 89 QGLTGSSDLRCFCYSGDVNAEPDSFAAVSLCGGLRGAFGYRGAEYVTSPLNPASAPAAQ 148
|||||
Db 13 QRLTGSSDLRCFCYSGDVNAEPDSFAAVSLCGGLRGAFGYRGAEYVTSPLNPASAPAAQ 72
QY 149 RNSQGAHLQRRGPGVGGSDPTSRGCVASGNWNPAILRALDPYKPRRTGSGESRRRSRG 208
|||||
Db 73 RNSQGAHLQRRGPGVGGSDPTSRGCVASGNWNPAILRALDPYKPRRTGSGESRRRSRG 132
QY 209 RAKRFVSIPIRYVETLVVADESVMKPHGADLEHYLLTLTAARLYRHPHSILNPINIVVK 268
|||||
Db 133 RAKRFVSIPIRYVETLVVADESVMKPHGADLEHYLLTLTAARLYRHPHSILNPINIVVK 192
QY 269 VILLRDRDSGPKVTGNAALTFRNFCAWOKKLNKVS DKHPEYWDTAILETRDLCGATCD 328
|||||
Db 193 VILLRDRDSGPKVTGNAALTFRNFCAWOKKLNKVS DKHPEYWDTAILETRDLCGATCD 252
QY 329 TLGMADVGTMDCKPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDMNKVCEEVFGKLRANH 388
|||||
Db 253 TLGMADVGTMDCKPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDMNKVCEEVFGKLRANH 312
QY 389 MMSPTLIQIDRANPWSACSAAITDLDGSGHCDLDDPSKPSISLPEDLPGASTYLSQC 448
|||||
Db 313 MMSPTLIQIDRANPWSACSAAITDLDGSGHCDLDDPSKPSISLPEDLPGASTYLSQC 372
QY 449 ELAFGVGSRPCPYMOCYTKLWCTGKAKGMVQCOTRHFPMADGTCGEGKCLKGACVERH 508
Db 373 ELAFGVGSRPCPYMOCYTKLWCTGKAKGMVQCOTRHFPMADGTCGEGKCLKGACVERH 432
QY 509 NLNKHVDSWAKWDYPGCSRTCGGVQLARR--QCTNPTPANGKCYCEGVVYKRSN 566
|||||
Db 433 NPKNYVDGPWAKWEYPGCSRTCGGGAQLARRVQATLPPLP-TGGKYCEGVVYKRSN 491
QY 567 LEPCPSSASGKSFR 580
|||||
Db 492 LEPCPSSASGKSFR 505

RESULT 6

AA49501
ID AAY49501 standard; Protein; 950 AA.

XX AC AAY49501;

XX DT 10-JAN-2000 (first entry)

XX DE Human METH1 protein.

XX KW Human; METH1; METH2; anti-angiogenic; metalloproteinase thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
XX arterial-venous malformation; immune deficiency.

XX OS Homo sapiens.

XX PN WO9937660-A1.

XX PD 29-JUL-1999. 99WO-0501313.
 XX PF 22-JAN-1999; 98US-0072298.
 XX PR 23-JAN-1998; 98US-0098539.
 XX PR 28-AUG-1998; 98US-0098539.
 XX PR (TRUE/) IRUELA-ARISPE L.
 XX PA (HAST/) HASTINGS G A.
 XX PA (RUBE/) RUBEN S M.
 XX PI IrueLa-Arispe L, Hastings GA, Ruben SM;
 XX WI: 1999-590684/50.
 XX DR N-PSDB; AA232000.
 XX PT New isolated metalloprotease thrombospondin polypeptides, useful for
 XX PT treating hyperproliferative disorders, cancers or autoimmune disorders
 XX PT
 XX ES Claim 10; Fig 1; 457pp; English.
 XX AA232000 and AA232001 encode, and AA49501 and AA49502 represent, human
 CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
 CC respectively. METH1 and METH2 have been found to be potent inhibitors of
 CC angiogenesis both in vitro and in vivo. They can be used for treating
 CC cancer and other disorders related to angiogenesis including abnormal
 CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
 CC endometrial bleeding disorders, diabetic retinopathy, some forms of
 CC macula degeneration, haemangiomas, and arterial-venous malformations.
 CC They may be useful in treating deficiencies or disorders of the immune
 CC system, by activating or inhibiting the proliferation, differentiation,
 CC or mobilization (chemotaxis) of immune cells. The etiology of these
 CC immune deficiencies or disorders may be genetic, somatic, such as
 CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
 CC toxins), or infectious. They can also be used to treat inflammatory
 CC conditions, both chronic and acute conditions. The products can also be
 CC used for detection and diagnosis. AA232002 to AA232080, and AA49503 to
 CC AA49511 represent sequences given in the exemplification of the present
 CC invention.
 XX Sequence 950 AA;
 XX Best Local Similarity 48.2%; Score 2486.5; DB 20; Length 950;
 XX Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;
 QY 1 MLLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPDGQGLIFQTAF 60
 DB 19 LLLLAALLAVSDALGRPSEDELVVP-ELE-----RAP---GHGTTLRLLHAF 64
 QY 61 QEDFVHLTPDAQTLAPAFSTHGLG-----VPLQGLTGGSSDLRRRCFYSGDVNAEPDSF 114
 DB 65 DQDLLELRPDSSFLAPGFTLQNVGRKSGSTPLP-----ETDLAHCFYSTVNGDPSA 119
 QY 115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLI0--RRGVPGG 165
 DB 120 AALSCEGVRGAFYLLGEAYFTIPLPAASERLATAAPCEKPPAPLQFHLLRRNQGDVGG 179
 QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRAGFCESRRRS 207
 DB 180 TCGVVDDEPRPTGKAETEDEGTEGEPQWS-----PQDPALQGVGP-TGTGS 230
 QY 208 GRAKRFVSIPIRVETLVVADESMVFGADLEHLLTLLATAARLYRHPSTLNPIVIVV 267
 DB 231 IRKKRFVSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPSTNSVSLVVV 290
 QY 268 KYLLLRDRDSGPKVTGNAAALTLRNFCAMOKLNKVKHPEYWDTAILFTRODLCGATTC 327
 DB 291 KILVTHDEQKGPVETNAALTLRNFCNQKHPNPPSDRDAEHYDTAILFTRODLCGSQTC 350
 QY 328 DTGLGMADVCTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDDAKQKASLVNVD 410
 DB 351 DTGLGMADVCTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDDAKQKASLVNVD 410
 QY 388 HMMSTPLTQIDRANPWSACSAAITLFDLSDGHCCLLDOPSKPISLPEDLPGASYTLSSQ 447
 DB 411 HMMASMLNLDHISQFSPSPCSAYMITSFLDNGHGECLMDKPNPQLPGLDLEPCTSYDANRQ 470
 QY 448 CELAFGVGSKPCP-YMOYCTKLWCTGKAKAGQVQCOTRHFPPWADGTSCGEGKLCAGACVE 506
 DB 471 COFTFGEDSKHCPDAASTCSTLWCTGTSGVLCVQTKHFPWADGTSCGEGKWCINGKCVN 530
 QY 507 RHNLNKH---RVDSGWAKWDYPGCSRTCCGGVQLARRCCTNPPTPANGKCYCEGVRYKYR 563
 DB 531 KTD-RKHFDTPPHGSMGMPGDCSRTCCGGVQVQTMRECDNPVPKNGKCYCEGVRYKYR 589
 QY 564 SCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSGVSPDRCKLCICRAN 623
 DB 590 SCNLEDCPDN-NGKTFREOCEAHNEFSKASGSPAVEMIPKYAGVSPKDRCKLICCAK 648
 QY 624 GTGYFVYLAQKVDGTLCSPDSTSVGVGKCIKAGCDNGLSKKRFKCGVCGGDNKSK 683
 DB 649 GIGYFVLPQKVVDGTPCSPDSTSVGVGQCVKAGCDRIIDSKKFKDKCGVCGGNGSTCK 708
 QY 684 KYTGFTKPMHGYNFVWALPAGASSIDIRQGYKGLIGDNYLALKNQSGKYLNGHEVW 743
 DB 709 KISGVSYSAPGYHDIITPTGATNIEVKQRNQRNNGSFLAIKAADGTYILNGDYTL 768
 QY 744 SAVERLVVKGSLRYSGTGTAVESLQASRPILPTVEVLSVGRMTPPRVYSFYLPEK 803
 DB 769 STLEQDINYKGVLYRSGSSAALIRSFPLKEPLTIQVLTGVNLRPKIKYTFVKKK 828
 QY 804 PREDKSSHPKDPGRGVSILNSVLSNQVEQDPPRPAWVAGSWGSPCSASCGSLQKRA 863
 DB 829 ----KES-----FNAIPF-----AWVTEEWGECSSKCELGWQRR 861
 QY 864 VDCRSAGQRTVPCADAAH--RPVETQAGE-PCPTWELSAWSPCSKSGRGFCRRSLKC 920
 DB 862 VECRDINGO---PASECAKEVPAPSTRPCADHPCPQWOLGEWSSCKTKGKGYKRRSLKC 918
 QY 921 VHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
 DB 919 LSHDGGVLSHESCDPLKPKHFDICTMAEC 949
 RESULT 7
 AAB73549
 ID AAB73549 standard; Protein: 950 AA.
 XX AAB73549;
 AC AAB73549;
 DT 07-AUG-2001 (first entry)
 XX Human ADAM-type metalloprotease MDT54, SEQ ID NO:4.
 DE Human: MDT54; ADAM-type metalloprotease; drug screening;
 KW A Disintegrin And Metalloprotease; cancer; arthritis.
 XX Homo sapiens.
 XX JP2001017183-A.
 XX 23-JAN-2001.
 XX 09-JUL-1999; 99JP-0196584.
 XX 09-JUL-1999; 99JP-0196584.
 PA (YAMA) YAMANOUCHI PHARM CO LTD.
 XX WPI: 2001-275950/29.
 DR N-PSDB; AAH20224.
 XX A new metal protease and its preparation for use as an anti-cancer and

PT anti-arthritis therapeutic -

PS Claim 1; Page 12-14; 22pp; Japanese.

CC The invention relates to the novel human ADAM (A Disintegrin And Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5 (AAB73550). The metalloproteases can be used for the treatment of cancers and arthritis. The invention also relates to the genes encoding MDTs4 and MDTs5, vectors and host cells containing the MDTs4 or MDTs5 genes, the recombinant production of MDTs4 and MDTs5, and antibody specific for MDTs4 or MDTs5, and methods of screening for compounds which modulate the activity of MDTs4 and/or MDTs5. The present sequence represents human MDTs4.

XX Sequence 950 AA;

Query Match 48.28; Score 2486.5; DB 22; Length 950;
Best Local Similarity 49.08; Pred. No. 7.7e-184;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

QY 1 MLLGILTLAFAGRTAGGSEPERVVPVIRLDPDINGRRYYWRGPEDSGQGLIFQITAF 60
DB 19 LLLAAALAVSDALGRSEDEELVVP-ELE-----RAP---GHGTRLRKHAF 64
QY 61 QEDFVHLTPDAQFLAPAFSTHGLG-----VPLQGLTGSSDLRCFYSCDVNAEPDSF 114
DB 65 DQGLDLRPPSSFLAPGFTIQNGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 119
QY 115 AAVSLGGLRGAFGYRGAEYISPLPNAS---APAAQRNSOGA---HLQ---RRGVPGG 165
DB 120 AALSCEGVGAFYLLGAEYITQPLPAASERLATAAPGEPAPAPLQPHLLKRNQGDVG 179
QY 166 PSG-----DPTSRC-----GVASG--WNPALRALDPKPRAPGFESESRERS 207
DB 180 TCGVVDDEPRFTGKAETDEDEGEDEGEQWNS-----PODPALQGVGP-TGTGS 230
QY 208 GRAKRFVSTPRVETLVYVADESVMKFGADLEHYLLTLLATAARLYRHPSLNPNIVVV 267
DB 231 IRKKRFVSSHRYVETMLVADQSMAEFGSLKHYLLTFLSVAARLYRHPSLRNSVSILVV 290
QY 268 KVLRLDRDSGPKVTGNAAALIRNFCAMQKLNKVSQKHPEYWDITLPTQDLGCATTC 327
DB 291 KILVTHDFQKGPETVSNAAALTRNFCNMQKHNPSPDRAEYDITLPTQDLGCGQTC 350
QY 328 DTGLMADYGTMDCKRRKSVTEDDGLPSAFTTAHELGHVFMNPHDNKVKVLEVEFGKLRAN 387
DB 351 DTGLMADYGTMDCKRRKSVTEDDGLPSAFTTAHELGHVFMNPHDPAQASLNGVQDS 410
QY 388 HMMSTPLTQIDRANPWSACSAITDPLDGHGDCCLDQPSKPLSLPELDPGASYTLISQQ 447
DB 411 HMMASMLNLDHSQSPWSPSCSAYMITSFLLDNGHGECLMDKPNPQLQPLDLPSTYDANRQ 470
QY 448 CELAFGVGSKPCP-YMQVCTKLWCTGKAKGMQVQTRFHWADJTGSGEGKGLKAGACVE 506
DB 471 CQFTGEDSKHCPDRAASTCSLWCTGSGVIVQCTKTFHWADJTGSGEGKGLKAGACVE 530
QY 507 RHNLNKH---RVDGSAWAKWDYPGCRTCGGGVQVLAARRQCFTNPANGAYCYCEGRVKYR 563
DB 531 KTD-RKHFTDPHGSWMGMPWGDSCRTCCGGGVQYTWRECDNPVPKNGKAYCYCEGRVKYR 589
QY 564 SCNLEPCSSASGFRFEOCEAFNGYHNSTNRLTAVAVVPKYSGVSPDKCKLICRAN 623
DB 590 SCNLEPCDPN-NKFTFREQCCAHNEFKASGGSPAVETIPKYAGVSPDKCKLICQAK 648
QY 624 GTGYFVLAPKVVDTGLSPDSTSVYGVQKICAGCDGNLGSKKRFPDQGVCGGDNKSK 683
DB 649 GIGYFVLQPKVVDGTPCSPDSTSVYGVQKICAGCDRIIDSKKFKDKCVCGGNGSTCK 708
QY 684 KYTGFTKPMHGYNFWAIPAGASSIDIRQYKGLIGLDNNYLAKNSQGYLLNGHFVV 743
DB 709 KISGSVTSKAPGYHDIITPTGATNIEYKQRNQSRNNGSFLAKADETYILNGDYTL 768
QY 744 SAVERDLVVKGSLLRYSGTGTAVESLOASRPILPLETLVEVLSVGKMTPPKRVRYFYLPE 803

DB 769 STLEQDIMYKGVLRYSGSSAALERIRSFPLKEPTIQVLTVGNALRPKIKYTFVKKK 828
QY 804 PREDKSSHPKDPGPFVSLHNSLSNQVQPDPPRPARVWAGSWGPCSCSGSLQKRA 863
DB 829 ---KES-----FNAITPFS-----AWVEWGECSKSCSLGWORRL 861
QY 864 VDCRSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCKSCSGRGFORSLKC 920
DB 862 VECRDINGQ---PASECAKEVKVPASTRPCADHPQWQLGEMWSSCKTCGKGYKRSKSLC 918
QY 921 VHGGRLLARDQCNLHRKPQBP-LDFCVLRPC 950
DB 919 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 949
RESULT 8
AAB50002
ID AAB50002 standard; Protein: 950 AA.
XX AC AAB50002;
XX DT 19-MAR-2001 (first entry)
XX DE Human METH1.
XX KW Human; METH1; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control.
XX OS Homo sapiens.
XX PN WO200071577-A1.
XX PD 30-NOV-2000.
XX PF 25-MAY-2000; 2000WO-US14462.
XX PR 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.
PR 22-FEB-2000; 2000US-0183792.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
PA (JONA/) JONAK Z L.
PA (TRUL/) TRULLI S H.
PA (FORN/) FORNWALD J A.
PA (TERR/) TERRETT J A.
XX IrueLa-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
PI Fornwald JA, Terrett JA;
XX WPI: 2001-025136/03.
DR N-PSDB; AAC90057.
XX METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT inhibit angiogenesis in the treatment of disorders such as cancer,
PT rheumatoid arthritis and psoriasis -

PS Claim 15; Fig 1; 768pp; English.

XX The present sequence is human METH1 (ME for metalloprotease and TH for thrombospondin). METH1 can be used for inhibiting angiogenesis in an individual, and for treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculogenesis, granulations, hypertrophic scars, nonunion fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, fibromuscular dysplasia, wound granulation, Crohn's disease or atherosclerosis. METH1 can also be used in birth control. METH1 can also be used in diagnostic methods for the prognosis of cancer.

XX Sequence 950 AA;

Query Match 48.2%; Score 2486.5; DB 22; Length 950;

Best Local Similarity 49.0%; Pred. No. 7.7e-184;

Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

Qy 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSJDOGLIFQITAF 60
Db LLLAALLAVSALGRPSEDEELVVP-ELE-----RAP--GHCFTRLRLHAF 64
Qy 61 QEDFLHLTPDAQPLAFSTFHLG-----VPLQGLTGGSDLRRCFYSGDVNAEPDSF 114
Db DQDLLELRDPSSFLAGFTLGVGRKSGSETPLP-----ETDLACFFSGTVNGDPSSA 119
Qy 115 AAVSLCGLRGAFGYRGAEEYVISPPLNPNAS---APAAORNSOGA---HLQ---RRGVPG 165
Db AALSICGVRGAFYLLGEAYFIQPLPAASERLATAPEKPPAPLOPHLLRRNRQDVG 179
Qy 166 PSG-----DPTSRC-----GVASG--WNPALLRALDPYKPRACFGESRRRS 207
Db TCGVVDDEPRPTGKAEDEDECTEGDEGPQMS-----PODPALQVGGP-TGTGS 230
Qy 208 GRAKRFVPIRYVETLVVADESVMKFGADLEHYLLTLLATAARLYRHPILNIPINIVV 267
Db IRKRFVSSHRYVETLVVAOSMAEFHSGGLKHYLLTFLSVAARLYKHPILRNSVSLVV 290
Qy 268 KVLRLDRSDGPKVTGNAALTLRNFCAWQKLNKVSDBEYVWOTAILFTQDLCAATTC 327
Db KILVTHDQKPEVTSNAALTLRNFCAWQKLNKVSDBEYVWOTAILFTQDLCAATTC 350
Qy 328 DTGLMADVGTCDPKRSCSVTDDGLSPAFTHAELGHVFNMPHDNVKVEEVFGKLRAN 387
Db DTGLMADVGTCDPKRSCSVTDDGLSPAFTHAELGHVFNMPHDNVKVEEVFGKLRAN 410
Qy 388 HMSPTLIQIDRANPWSACSAIITDFLDSHGDCLLDQPSKPISLPEIIFGASYTLISQ 447
Db HMASMLNLDHSPWSPSCSAIITDFLDSHGDCLLDQPSKPISLPEIIFGASYTLISQ 470
Qy 448 CELAFGVSKPCP-YMOYCTKLWCTGAKGQWVOTRHPWADGTSGEKLCLKAGACVE 506
Db COFTFGEDSKHCPDAASTCTSLWCTGSGVLVVCOTRHPWADGTSGEKLCLKAGACVE 530
Qy 507 RHNLNKH---RVDGSKAWKDPYPCSRCTGGGVQALRQCTNPTPANGKYCEGVRYKYR 563
Db KTD-RKHEDTFHSGWMGWPWDCSRCTGGGVQALRQCTNPTPANGKYCEGVRYKYR 589
Qy 564 SCNLEPCSSAGSKFRFEEQCEANFYNGHSTNRLTLAVAWPKYSGVSPDKCKLICRAN 623
Db SCNLEPCSSAGSKFRFEEQCEANFYNGHSTNRLTLAVAWPKYSGVSPDKCKLICRAN 648
Qy 624 GTGYEYVLAPKVDGTLCSPTDSCVCGKICAKGDCNLSGSKKRDCKCGCGDNKSK 683
Db GTGYEYVLAPKVDGTLCSPTDSCVCGKICAKGDCNLSGSKKRDCKCGCGDNKSK 708
Qy 684 KVTGLTFPMHGYNFWVAIPACASSIDIRQYKGLIGDDNYLALKNQSKYLINGHFVV 743
Db KVTGLTFPMHGYNFWVAIPACASSIDIRQYKGLIGDDNYLALKNQSKYLINGHFVV 768

Qy 744 SAVERDLVVKSGLLRYSGTGTAVESLQASRILEPLVEVLVSVCKMTPPRRVRSFYLPKE 803
Db STLEQDIMYKGVWLRYSGSSAALERINSFPLKEPLTQVLTGVGNALRPKIKTYFVYKK 828
Qy 804 PREKSSHPKDPGRPSVLHRSVLSLNSQVEQPDPRPARVAGSWGPCSCGSLQKRA 863
Db ---KES-----FNAIPTFS-----AWVIEWGECSSKCELGWQRL 861
Qy 864 VDCRSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCKSCGSGFQRRSLK 920
Db VECDINGQ---PASECAKEVYPASTRCPADHPCPQOLGEMSSCSTCGKGYKKRSLK 918
Qy 921 VGHGRLIARDQCNLHRRKQPB-LDFCVLRPC 950
Db LSHDGVLSHESCDPLKKPKHFIDFCTMAEC 949

RESULT 9

AAW80285

ID AAW80285 standard; Protein; 967 AA.

AC AAW80285;

DT 19-JAN-1999 (first entry)

XX Human integrin ligand polypeptide ITGL-TSP.

XX ITGL-TSP; integrin ligand; angiogenic disease; cancer; atherosclerosis;
KW chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;
KW macular degeneration; diabetic retinopathy; Alzheimer's disease; human;
XX restenosis.

OS Homo sapiens.

PN EP874050-A2.

PD 28-OCT-1998.

PF 27-JAN-1998; 98EP-0300575.

PR 24-APR-1997; 97US-0845496.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (SMK) SMITHKLINE BEECHAM CORP.

PA (SMK) SMITHKLINE BEECHAM PLC.

XX Fronwald JA, Hastings GA, Jonak ZL, Terrett JA;

PI Trulli SH;

XX WPI: 1998-544643/47.

DR N-PSDB; AAV66508.

XX DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat
XX angiogenic diseases, restenosis, Alzheimer's disease and in tissue
XX remodeling

PS Claim 11; Pages 6-9; 24pp; English.

XX This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP
XX polypeptides can be used in the treatment of angiogenic diseases such as
XX cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid
XX arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,
XX restenosis, Alzheimer's disease and tissue remodeling. They can be used
XX to treat a subject in need of enhanced activity or expression of the
XX ITGL-TSP polypeptide.

XX Sequence 967 AA;

Query Match 48.2%; Score 2486.5; DB 19; Length 967;

Best Local Similarity 49.0%; Pred. No. 7.9e-184;

Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

Qy 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSJDOGLIFQITAF 60

```
Db 36 LLLLAALLAVSDALGRSEDEELVVP-ELE-----RAP-----GCTTTLRLHAF 81
QY 61 QEDFVHLTPDAQFLAPAFSTHELG-----VPLQGLTGGSSDLRPFYSGDVNAEPDSF 114
Db 82 DQDLLELRPOSSFAPGFTLIONVGRKSGSETPLP-----EIDLAHGFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFYGRGAEYVISLPLNAS-----APAAORNSOGA-----HLIQ--RRGVPGG 165
Db 137 AALSCEGVRGAFYLLGAEYFIQPLPAASERLATAAAGEXPAPLQHLRLRNQGVGG 196
QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFEGESRRRS 207
Db 197 TCGVVDDPRPTGKAETDEDEGTEGEGPQWS-----PQDPAQLQVQGP-TGTGS 247
QY 208 GRKRFVSIPRYVETLVVVADESVMKFGADLEHYLLTLLATAARLYRHPISILNPINVVV 267
Db 248 LRKRFVSSHRYVETLMVADGMAEFHSGLLGKHYLLTFVSAARLYRHPISILNPINVVV 307
QY 268 KVLRLDRDSGPKYVTGNAALTNRFCQAKWKLKVSQKHPYWDITLFTKQDLGATTC 327
Db 308 KILVHIDBOKPEVTSNAALTNRFCQAKWKLKVSQKHPYWDITLFTKQDLGATTC 367
QY 328 DTLGADVGTWCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVYVVEEVFGKLRAN 387
Db 368 DTLGADVGTWCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVYVVEEVFGKLRAN 427
QY 388 HMSFTLIQIDRANPWSACSAAITLDFDLSGHGCLLDQSKPISLPEDLPAGASYTLISQ 447
Db 428 HMASLNLNDHSPWSPCSAYMTSFLDNGHGECLMDKFNQPLQPLGDLPGTSYDANRQ 487
QY 448 CELAFGVSGRCP- YMOYTKLWCTGKAKGMVQCTRHFPWADGISGEGSKLCLKGACVE 506
Db 488 COFTFGEDSKHCPDAASTCTLTGTSGGLVLCQTKHFPWADGISGEGSKLCLKGACVE 547
QY 507 RHNLNKH---RVGDSWAKWDPYGCSTCGGVLQARQCTNTPANGAYCGRVRYR 563
Db 548 KTD-RKEDTPFHGSGWGMGPMWDCSRTCGGVOYVTRNCDNPNVNGGKYGCGKRVYR 606
QY 564 SCNLEPCPSAGSKSFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCLICRAN 623
Db 607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPADRCCLKCQAK 665
QY 624 GTGYFVYLAQVVDGTLCSFDSSTVSVQGCITKAGDGNLGSKKRFDKCVCGGDNKSK 683
Db 666 GIGYFFVLQPKVVDGTFPCSPDSTVSVQGCITKAGDGNLGSKKRFDKCVCGGDNKSK 725
QY 684 KVTGLFTKPMHGNFVVAIPAGASSIDIRORYKGLIGDNNYLAKNSQKYLINGHFVV 743
Db 726 KISGVSYSAPGYHDIITIGATNIEVKORNGRNGNSFLAIKAADSTYILNGDYTL 785
QY 744 SAVERDLVWGLSLRYSGTAVESLOASRPILPTVEVLSYGRKMTPPFVYSFYLPKE 803
Db 786 STLEQDINMYGVVLYRYSGSAALERTSRFSPLKEPLTIOVLTVGNALRPKIKYTFVKK 845
QY 804 PREDKSHHPDRPGSVLHNSVLSLNOVQPDPRPARVAGSWGPCSAAGSGGLQKRA 863
Db 846 ----KES-----FNAITFS-----AWVIEEWECSSCELGWQORL 878
QY 864 VDCRGSAGQRTVPACDAAH--RVETQACGE-PCPTWELSAWSPCSKSGRGFORSLAC 920
Db 879 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWQGEWSSCKTGKGGKRSKLC 935
QY 921 VGHGGRLLARDQCNLHRKQGE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKPKHFDICTMAEC 966
```

RESULT 10

AAB50011

ID AAB50011 standard; Protein; 968 AA.

XX

AC AAB50011;

```
XX DT 19-MAR-2001 (first entry)
XX DE Protein; SEQ ID 125.
XX KW Human; METH1; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Weber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control.
XX OS Homo sapiens.
XX PN WO200071577-A1.
XX PD 30-NOV-2000.
XX PF 25-MAY-2000; 2000WO-US14462.
XX PR 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.
PR 22-FEB-2000; 2000US-0183792.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (SMK) SMITHKLINE BEECHAM CORP.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (IRUE/) IREUA-ARISPE L.
PA (RUBE/) RUBEN S M.
PA (JONA/) JONAK Z L.
PA (TRUL/) TRULLI S H.
PA (FORN/) FORNWALD J A.
PA (TERR/) TERRETT J A.
XX Iruela-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
PI Fornwald JA, Terrett JA;
XX WPI; 2001-025136/03.
XX METH1 and METH2 polynucleotides and encoded polypeptides, used to
inhibit angiogenesis in the treatment of disorders such as cancer,
rheumatoid arthritis and psoriasis.
XX Claim 15; Pages 759-763; 768pp; English.
CC The present invention relates to human METH1 and METH2 (ME for
metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
METH can be used for inhibiting angiogenesis in an individual, and for
treating cancer, benign tumours, an ocular angiogenic disease,
rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
coronary collaterals, cerebral collaterals, arteriovenous malformations,
ischaemic limb angiogenesis, Osler-Weber syndrome, plaque
neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
fibromuscular dysplasia, wound granulation, Crohn's disease or
atherosclerosis. METH can also be used in birth control. METH can also
be used in diagnostic methods for the prognosis of cancer. The present
sequence is a protein isolated in the present invention.
```

Sequence 968 AA;

Query Match

48.2%; Score 2486.5; DB 22; Length 968;

Best Local Similarity 49.0%; Pred. No. 7.9e-184;

Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

```
QY 1 MLLGLILTLAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPEDSGCGQLIFQITAF 60
DB 37 LLLAAALLAVSDALGRPEDEELVVP-ELE-----RAP---GHGTTLRLLHAF 82
QY 61 QEDFYHLPTDPAQFLAPAFSTHGLG-----VPLQGLTGSSDLRRRCFYKDVNAEPDSF 114
DB 83 DQDLLELRPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCYFSGTVNGDPSSA 137
QY 115 AAVSLCGGLRGAFYGAEGYVISPPLNAS---APAAQRNSOGA-----HLLQ---RRGVPG 165
DB 138 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLRRNQGDVGG 197
QY 166 PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRRAGFGESESRRS 207
DB 198 TCGVVDDEPRPTGKAETDEDETEGEDEGPWS-----PQPALQGVGP-TGTGS 248
QY 208 GRAKREVSIPRYVETLVVADESVMKFGADLEHYLLTLATAARLYRHPISILNPINIVV 267
DB 249 IRKREVSSSHRYVETMLVADQSMAEFHGSLKHYLLTLFVAARLYKHPISIRNSVSLVV 308
QY 268 KVLRLDRDGGPKVTGNAALTLENFCAWOKKLNKYSKHPYWDTAILFTRODLGATTC 327
DB 309 KILVIHQKPEVTGNAALTLENFCAWOKKLNKYSKHPYWDTAILFTRODLGATTC 368
QY 328 DTLCMADVTGMDCKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVKNVHEVFGKLRAN 387
DB 369 DTLCMADVTGMDCKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVKNVHEVFGKLRAN 428
QY 388 HMASLNLSDHQPWSPCSAYMITSLFNGHGECLMDKPNP IQLPGDLPSTSYDANRQ 447
DB 429 HMASLNLSDHQPWSPCSAYMITSLFNGHGECLMDKPNP IQLPGDLPSTSYDANRQ 488
QY 448 CELAFGVGSKPCP-YMQYCTKLAQCTKAKGOMVCOQTHPEPMDGTSCGEGKCLKACAYE 506
DB 489 COFTFGEDSKHCDPAASTCTSLMCTGTSGGVVQCQTHPEPMDGTSCGEGKCLKACAYE 548
QY 507 RHNLNKH---RVDGSWAKMDPYGPCSRTCGGVQGLARRQCTNPTPANGKYGCEGVKRYR 563
DB 549 KTD-RKHEDFPFHGSMGMWGPWDCSRTCGGVQGLARRQCTNPTPANGKYGCEGVKRYR 607
QY 564 SCNLEPCPSASGSKSFPREOCEAFNGYHNSTNLTAVANVPKYSVSPEDCKLTCRAN 623
DB 608 SCNLEPCPSASGSKSFPREOCEAFNGYHNSTNLTAVANVPKYSVSPEDCKLTCRAN 666
QY 624 GTGYFVYLAPKVDGTLCSPDSTSVCOGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSK 683
DB 667 GIGYFVYLAPKVDGTLCSPDSTSVCOGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSK 726
QY 684 KVTGLTKPMHGNVFNVAIPAGASSIDIRORYKGLIGDNDYLNALKNQSKYLLNGHFVV 743
DB 727 KISGVTSAKPGYHDIITPTGATNIEVKQNRGSRNNGSFLAKAADGYTLNGDYTL 786
QY 744 SAVERDLVWGLLRYSGTGTAVESLQASRPILPTEVLSVGMKTPPRVRYSFYLPKE 803
DB 787 STLQDLMYKGVLLRYSGSSAALIRISFPLKEPLTIOVLTVGNALRPAIKTYFVKK 846
QY 804 PREKSHPKDPRGPSVLHNSVLSNQVEQDPRPARWAGSWGPCSCSGSLQKRA 863
DB 847 -----KES-----FNAIPFS-----AWVIEWEGCSKSCELGWORRL 879
QY 864 VDCRGSAGORTVPACDAH--RPVETOACGE-PCPTWELSAWPCSKSCGRGQRSLKC 920
DB 880 VECRDINGO---PASECAKEKVPASTRPCADHPQPMOLGENWSCSKTCGKYKRSKLC 936
QY 921 VGHGRLARDQCNLHRRKPOE-LDFCVLRPC 950
DB 937 LSHDGVLSHESCDPLKPKPHFDICTMAEC 967
```

RESULT 11

AA04142

ID AA04142 standard; Protein; 967 AA.

```
XX AC AA04142;
XX AC 15-JUN-1999 (first entry)
XX DE Human Tango-71 protein.
XX KW Human; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis;
XX KW detection.
XX OS Homo sapiens.
XX PN WO9907850-A1.
XX PD 18-FEB-1999.
XX PF 06-AUG-1998; 98WO-US16502.
XX PR 05-SEP-1997; 97US-0058108.
XX PR 06-AUG-1997; 97US-0054966.
XX PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX PI Goodearl ADJ, Holtzman DA;
XX DR WPI; 1999-167426/14.
XX DR N-PSDB; AAX19955.
XX PT New TANGO polypeptides and nucleic acids encoding them - useful as
XX PT diagnostic agents and for treating disorders caused by aberrant
XX PS expression of TANGO
XX PS Claim 8; Fig 1; 84pp; English.
XX CC The present sequence represents human Tango-71. Tango polypeptides are
XX CC useful for identifying compounds which bind the polypeptide via direct
XX CC binding, competition binding assays or Tango-71, -73, -74, 76 or -83-
XX CC mediated signal transduction. Tango polypeptides are also useful for
XX CC identifying modulating compounds by determining effect on Tango activity.
XX CC Tango polypeptides and nucleic acids are useful for diagnosing diseases
XX CC related to aberrant expression of Tango, and Tango polypeptides are
XX CC useful for raising antibodies which can be used in diagnostic assays for
XX CC detection of Tango, and also for generating anti-idiotypic antibodies for
XX CC prevention and protection.
XX SQ Sequence 967 AA;
```

Query Match 48.1%; Score 2485.5; DB 20; Length 967;
Best Local Similarity 49.0%; Pred. No. 9.5e-184;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

```
QY 1 MLLGLILTLAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPEDSGCGQLIFQITAF 60
DB 36 LLLAAALLAVSDALGRPEDEELVVP-ELE-----RAP---GHGTTLRLLHAF 81
QY 61 QEDFYHLPTDPAQFLAPAFSTHGLG-----VPLQGLTGSSDLRRRCFYSGDVNAEPDSF 114
DB 82 DQDLLELRPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCYFSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFYGAEGYVISPPLNAS---APAAQRNSOGA-----HLLQ---RRGVPG 165
DB 137 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLRRNQGDVGG 196
QY 166 PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRRAGFGESESRRS 207
DB 197 TCGVVDDEPRPTGKAETDEDETEGEDEGPWS-----PQPALQGVGP-TGTGS 247
QY 208 GRAKREVSIPRYVETLVVADESVMKFGADLEHYLLTLATAARLYRHPISILNPINIVV 267
DB 248 IRKREVSSSHRYVETMLVADQSMAEFHGSLKHYLLTLFVAARLYKHPISIRNSVSLVV 307
QY 268 KVLRLDRDGGPKVTGNAALTLENFCAWOKKLNKYSKHPYWDTAILFTRODLGATTC 327
DB 309 KILVIHQKPEVTGNAALTLENFCAWOKKLNKYSKHPYWDTAILFTRODLGATTC 368
```

Db 308 KILVTHDQKGPVTSNAALTRNCWNQKOHNPSPDRDAEHYDTAILFTIHDLCGSGTC 367
 QY 328 DTGLMADVTWCDPKRCSVTEDDGLPSAFTTAHELGHVFNPHDNDVNVKVLVEVEGKLRAN 387
 Db 368 DTGLMADVTWCDPKRCSVTEDDGLQNAFTAHELGHVFNPHDDAKQACASLNGVQDS 427
 QY 388 HMMSTLTIQIDRANPWSACSAIIITDILDSHGDCILDDQPSKPLSLPEDI PGASYTTLSSQ 447
 Db 428 HMMASMLSNLDHSPWSPSCSAYMITSLFNDHGCECLMDKPNQPIQLPCDLPGTISYDANRQ 487
 QY 448 CELAFGVGSKPCP-YMOYCTKLWCTGKAKGWVCOTRHFPAWADTSCGEGKCLKLGACVE 506
 Db 488 COFTFGEDSKKCPDAASTCTSLWCTGSGGVLCVOTRHFPAWADTSCGEGKWCINGKCVN 547
 QY 507 RHNLNKH---RVDSWAKNDPYGCSRTCGGGVQLARRQCTNPANGKAYCEGVRYK 563
 Db 548 KTD-RKHEDTTFHSGWGMWPGDCSRTCGGGVYTHREDNPVPKNGGYCEGKRVYR 606
 QY 564 SCNLEPCSSASGKSFREEQCEAFNGYNHSTNRUTLAVAVVPKYSGVSPKDKCKLICRAN 623
 Db 607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKAVGSPKDRCKLICQAK 665
 QY 624 GTGYFYVLAPKVDGTCLSPDSTVCGQKCIKAGCDENLGSKKRPDKQVCGGDNKSK 683
 Db 666 GIGYFFVLQPKVDGTCPSPDSTVCGQGVKACCDRIIDSKKKPKDKCVCGGNGSTCK 725
 QY 684 KVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDIDNYLALKNSQSKYLLINGHFV 743
 Db 726 KISGVSISAKPGYHDIITPTGATNIEVKQNRQNSGFLAIRKAADITYILNGDYTL 785
 QY 744 SAVERDLVWKSLLRYSGTAVESLQASRPLEPLTEVLSVCKMTPPPVRYSYFLPK 803
 Db 786 STLEQDLMYKGLVRLYSGSAAALIRSFSPLEPLTIQVLTCGNALRPKIKYTFYVKK 845
 QY 804 PREDKSHPKDPRGSPVLSNLSNOVEQPDPRPARVWAGWGPCSACSGSLQKRA 863
 Db 846 ---KES-----FNAITFS-----AWIIEWGECSSKSELGWORL 878
 QY 864 VDCRSGAGRTVPACDAH--RPVETQACE-PCPTWELSAWSPKSCSGKGFQKSLKC 920
 Db 879 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWOLGEMWSSCKTCGKGYKKSLLKC 935
 QY 921 VGHGRLARDQCNLHKKPQE-LDFCVLRPC 950
 Db 936 LSHDGGVLSHESCDPLKKPKHFIIDFCTMAEC 966

RESULT 12
 AAW78189
 ID AAW78189 standard; Protein; 967 AA.
 XX
 AC AAW78189;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene 64 clone H0UCQ17.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 40
 FT /label= unknown
 FT Misc-difference 45
 FT /label= unknown

FT Misc-difference 169 /label= unknown
 FT Misc-difference 293 /label= unknown
 FT Misc-difference 297 /label= unknown
 FT Misc-difference 557 /label= unknown
 XX
 PN W09856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 11-JUN-1998; 98WO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058668.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 02-OCT-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 XX
 WPI; 1999-080881/07.
 N-PSDB; AAX04374.
 XX
 New isolated human genes and the secreted polypeptides they encode -
 disorders, immune diseases, inflammation or blood disorders
 XX
 Claim 11; Page 297-300; 380pp; English.
 XX
 This sequence represents a secreted human protein encoded by the gene
 clone detailed in the descriptor line. The gene can be used to generate
 fusion proteins by linking to the gene to a human immunoglobulin FC
 portion (e.g. AAX04302) for increasing the stability of the fused
 protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).

XX	Sequence	967 AA;	
SQ	Query Match	47.7%; Score 2464.5; DB 20; Length 967;	
	Best Local Similarity	48.6%; Pred. No. 4e-182;	
	Matches	482; Conservative 154; Mismatches 254; Indels 101; Gaps 24;	
QY	1	MLLLGTLTAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSCVQGLIFQITAF	60
DB	36	LLLLXALXVSDALGRPEDEELVVP-ELE-----RAP---GHGTRLELHAF	81
QY	61	QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGSSDRLRRCFYSGDVNAEPDSF	114
DB	82	DOQDLLELPDSFLAPGTLQNVGRKSGSEPLP-----ETDLAHCFFYSGTVNGDPSSA	136
QY	115	AAVSLCGGLRGAFGYRGAEEYVISPPLNAS---APAAQRNSQGA---HLLQ---RRGVGG	165
DB	137	AALSLCEGVRFAYLLGEAYFIQPLPAASERLXTAAGPEKPPAPLOFHILRRNQDVG	196
QY	166	PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRASIGESRRRS	207
DB	197	TCGVWDEPRPTCKATEDEDETEDEGPQWS-----PDPALGVGP-TGTGS	247
QY	208	GRAKRFVSPRYVETLLVVADESVMKPHGADLEHYLLTLTAAKLYRHPSETLNPNIVVV	267
DB	248	IRKRFVSSHRYVETMLVADQSMAPFHGSLKHYLLTLFSVAARKKHPXLRNSVSLVV	307
QY	268	KVLLLRDSDGPKVTGNAALTNRFCWAKKLNKVKSDKHPEYWDIAILLTIRODLCGATTC	327
DB	308	KILVIHQEQGPEVTGNAALTNRFCWAKKLNKVKSDKHPEYWDIAILLTIRODLCGATTC	367
QY	328	DTLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVKNVCEVFGKLRAN	387
DB	368	DTLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVKNVCEVFGKLRAN	427
QY	388	HMSPTLIQIDRANPSACSAALITDFLSGHGCDLLDQSPISLPEDLPAGASYTSLSQ	447
DB	428	HMASMLNLDHQPWSPSCSAYMITFSLDNGHCELDMPQNPDIQPLGDLPGTSDYDANQ	487
QY	448	CELAFGVSGPCP-YMOCYTKLWCKAKGMVQCOTRHPFADPSTGCPKCLKAGACVE	506
DB	488	COFTFGEDSKHCPDAASTCTSLWCTGTSGVLVVCQTHKFPWADSTSGCEKNCINGKCVX	547
QY	507	RHNLNKH---RVGDSNAKWDYPGCSRTCGGVQVLARQCTNTPANGAGYCEGVRYKR	563
DB	548	KTD-RKHFDPTFHGSGWMPGWDSCSRTCGGVQVYTMRECDNVPKNGGKYCEGVRYR	606
QY	564	SCNLEPCSSASGSKFREEOCAFNGYNHSTNRLTLAVAVPKYSGVSPDKCKLICRAN	623
DB	607	SCNLEPCPDN-NGKTFREEOCAHNEFSKASFGSGPAVEIPKYAGVSPADCKLICQAK	665
QY	624	GTGYFVYVLPKVDGTLCSDPSTSVCCQKCIKAGCDNGLSKKRFDKCFVCGGDNKSCK	683
DB	666	GIGYFVYVLPKVDGTPCSPDSTSVCCQKCIKAGCDNGLSKKRFDKCFVCGGDNKSCK	725
QY	684	KVTGLFTKPMHGYNFVAIPAGASSIDIRORYKGLIGDDNYLALNKSQKYLINGHVV	743
DB	726	KISGVTSAPKGYHDITITPTGATNIEVKQRNQRNNGSFLAIKAADGYIILNDYTL	785
QY	744	SAVERDLVWKSLLRYSGTGTAVESLOASRPILPELTVESLVGKMTPTPVRYFSYLPKE	803
DB	786	STLEQDIWKGVVLYRSGSSNALERISFSLKEPLTQVLTGNALRPAIKTYFYVKKK	845
QY	804	PREDKSHHPKDPGRPSVLSLNSVLSNOVEQDDRRPARVAGSWGPCSSACSGGLQKRA	863
DB	846	-----KES-----FNAIPTFS-----ANVIEEWGCSKSCGLGWQRL	878
QY	864	VDCRGSAGQTVACDAAH---RPVETOACGE-PCPTWELSNWSPCSKSGRGQRRLSKC	920
DB	879	VECDRIQO---PASECAKEVKPASTRPCADHPCPQOMOLGEWSSCKTCGKGYAKRLK	935
QY	921	VGHGRLIARDQCNLHRKPOB-LDFCVLRPC	950

Db 936 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966

RESULT 13

AA53899

ID AAY53899 standard; Protein; 950 AA.

XX AAY53899;

XX 13-MAR-2000 (first entry)

XX Amino acid sequence of a murine ADAMTS-1 protein.

XX GON-1; metalloproteinase; cell migration; modulator; ADAMTS-1;

XX metalloproteinase domain; thrombospondin domain; abnormal cell migration;

XX organ shaping; sterility; cancer metastasis.

XX Mus sp.

XX WO9961656-A2.

XX 02-DEC-1999.

XX 28-MAY-1999; 99WO-US11918.

XX 29-MAY-1998; 98US-0087170.

XX 13-APR-1999; 99US-0129023.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

XX Kimble JE, Belloch RH;

XX WPI; 2000-072633/06.

XX Identifying modulators of proteins containing metalloprotease and

XX thrombospondin domains, potentially useful for controlling cell

XX migration and organ shaping

XX Disclosure; Fig 1C; 60pp; English.

XX The present sequence represents a murine ADAMTS-1 protein. ADAMTS-1 is

XX metalloproteinase. The specification describes another related

XX metalloproteinase, a C. elegans GON-1 protein, that lacks a transmembrane

XX domain and possesses a predicted metalloprotease domain between residues

XX 269-456. In C. elegans hermaphrodites, GON-1 is required for migration of

XX two distal tip cells to produce elongated tubes, whereas in males, GON-1

XX is required for migration of a single linker cell to produce a single

XX elongated tube. The protein is used in the method of the invention. The

XX specification describes a method for identifying a modulator of a

XX protein that contains a metalloproteinase domain and a thrombospondin

XX domain. The method comprises treating a target organism, having a

XX developing gonadal cell that is responsive to the protein, with a test

XX compound, and determining any change in migration or shape of the cell

XX attributable to the test compound. The compounds identified are

XX potential therapeutic modulators of abnormal cell migration and organ

XX shaping, e.g. for rendering animals (specifically nematodes) sterile

XX and for inhibiting cancer metastases.

XX SQ Sequence 950 AA;

Query Match 47.4%; Score 2448.5; DB 21; Length 950;

Best Local Similarity 47.7%; Pred. No. 6.9e-181;

Matches 476; Conservative 159; Mismatches 249; Indels 113; Gaps 24;

QY 1 MLLGLTLTAFAGRTAGG--SEPEREVVPIRLDPDINGRRYYWRGP-EDSCDGLIFQI 57

DB 20 LLLLSITMLLCAGAHGRPTDEELVLP-SLE-----RAPGHDSTTTL--RL 66

QY 58 TAFQEDFYHLTPDAQFLAPAFSTEHLGVLPLOLTGGS-----SDLRRCFFYSGIWNNA 109

DB 67 DAFGQQLHLKLPDSCDGLFAPGTLQTV-----GRSPGSEAHQHLDTGLAHCFFYSGTVNG 121

QY 110 EPDSFAAVSLCGGLRGAFGYRGAEEYVISPPLNAS-----APAAQRNSQGA----HLLQRRG 161

Db 122 DPGSAALSLCGRVGAETVLOGEEFFIQAPGVATERLAPAVPEESSARQFHILRR- 180
 QY 162 VPGSGDPTSCGVASGWNPAIRLALDPYKPRRAGFESRRR----- 206
 Db 181 -----RRSGGAKCGVMD-----DETLPSTDSRPESQNTNQMVFDPDTPQDAGK 226
 QY 207 -----SGRAKRFVSPRYVETLVVADESVMKFGADLEHYLLTLLATAARVHRPISILNPI 262
 Db 227 SGPGSIRKRRVSSPRYVETLVVADQSMADFGSLGKHYLLTLFVAAEFYKHFSIRNSI 286
 QY 263 NIVVVKVLLLRDRSGPKVTGNAALTNRNFCAMQKLNKVDKHEPYMDIAILFTRQDLC 322
 Db 287 SLVVVKILVIYEEQKGPVTSNAALTNRNFCAMQKLNKVDKHEPYMDIAILFTRQDLC 346
 QY 323 GATTCDTLGMADVGTMCDPKSCSVIEDDGLPSAFTTAHELGHVFNPHNVKVCVEVFG 382
 Db 347 GSHTCDTLGMADVGTMCDPKSCSVIEDDGLPSAFTTAHELGHVFNPHNVKVCVEVFG 405
 QY 383 KLRANHMSPITLIQIDRANPWSACSAALITDLSHGDCILLDQPSKPISELPDPCASY 442
 Db 406 VSGSHLMASMLSLDHSQWSPSCSAYMVTSLDNGHGCLMDKPQNPXIPSDLPGLTY 465
 QY 443 TLSQOCELAFVGGKPCP-YMQYCTKLCWGTGKAGQWVCQTRHPFPNADGHSCEGKLC 501
 Db 466 DANROCCFTFGESKHCPCDAASTCTTLCWGTGKAGQWVCQTRHPFPNADGHSCEGKLC 525
 QY 502 GACVERHNLNK-----RVDGSAKWDYPGCSRTCCGGVQLARQCTNPTFANGKICEGV 558
 Db 526 GKCVNKTDM-KHFTATPVHSGWPGWGDSCRTCCGGVQVTMRCDNPVEKNGGKCEG 584
 QY 559 RVKVRONLEPCPSASGKSFREBQCEAFNGYNHSTNRLTLAVAWVFYKSGVSPDKCKL 618
 Db 585 RVRVRSNIEDCPDN-NGKTFREBQCEAHNEFSKASFGNEPTVETPKYAGVSPDKCKL 643
 QY 619 ICRANGYGVVLPKVDGTLSPDSTVCGVQKCIKACDGNLGSKKTDKCGVCGGD 678
 Db 644 TCEAKGIGYFVLQPKVVDGTLSPDSTVCGVQKCIKACDGNLGSKKTDKCGVCGGD 703
 QY 679 NKSCKVTGLTFPMHGYNEVVAIPAGASSIDIRQRYKELIGDNYLALKNQSGYLLN 738
 Db 704 GSTCKMKGIVTSRPGYHDIVTIPAGATNIEVKHNRQSRNNGSEFLAIRADGYILN 763
 QY 739 GHEVVSVERDLVVKSLRYSGTGAVESLQASRPILPTVEVLSVGKMTTPRVRYSF 798
 Db 764 GNFTLSTLEQDLYKGVLRYSGSSAALIRISFSLKEPLTIQVLMVGHALRPKIKFTY 823
 QY 799 YLPKEPREDKSSHDKPRGSPVLHNSVLSLNOVEQDPPARWVAGSWGPCASCGSG 858
 Db 824 FMKKKTES-----FNAIPTFS-----EWVIEENGESKTCGSG 856
 QY 859 LQKRVDCRGSAGRTVPACDAAH--RPVETQACG-PCTWELSAWSPCKSKCGRGQR 915
 Db 857 WQRRVVOCRDINGH--PASECAKEKVPASTRPCADLPCHMQVQWDSWPCSKTCGKYKK 913
 QY 916 R-SLKCVGHGGRLLARDONLHKPQES-LDFCVLRFC 950
 Db 914 RCLKCVSHGGVLSNESCDPLAKPKHYIDFCTLTCQ 950
 RESULT 14
 AAB21265
 ID AAB21265 standard; Protein; 896 AA.
 AC AAB21265;
 XX
 XX 23-FEB-2001 (first entry)
 XX Mouse metalloproteinase ADAMTS-1.
 XX Mouse; ADAMTS-1; metalloproteinase; ADAM;
 KW a disintegrin and metalloproteinase domain; thrombospondin domain;
 KW vaccine; neutropic; neuroprotective; antiparkinsonian;

KW cerebroprotective; cytostatic; antiarthritic; immunosuppressive;
 KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;
 XX autoimmune disease; brain tumour; brain injury.
 OS Mus musculus.
 XX W0200053774-A2.
 PN 14-SEP-2000.
 XX 08-MAR-2000; 2000WO-US06237.
 PF 08-MAR-1999; 99US-0264585.
 PR (NEUR-) NEUROCRINE BIOSCIENCES INC.
 XX Kelner GS, Clark M, Maki RA;
 XX WPI; 2000-594326/56.
 DR Polynucleotide encoding novel members of a disintegrin,
 XX metalloproteinase and thrombospondin domain protein family used to
 PT prevent and treat Alzheimer's disease, cancer and autoimmune diseases -
 PS Disclosure; Fig 17; 129pp; English.
 XX The present sequence is mouse metalloproteinase ADAMTS-1. The
 CC ADAMTS family of proteins is closely related to the ADAM (A Disintegrin
 CC and Metalloproteinase Domain) family. Members of the ADAMTS family
 CC contain a thrombospondin domain in addition to the disintegrin and
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are
 CC useful for the manufacture of medicaments for treating conditions
 CC associated with neuroinflammation and/or neurodegeneration, such as
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also
 CC useful for treating conditions associated with cell proliferation, cell
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis
 CC and autoimmune diseases. They can be used to treat patients afflicted
 CC with an invasive tumour, a brain tumour or brain injury.
 XX Sequence 896 AA;

Query Match 45.0%; Score 2325; DB 21; Length 896;
 Best Local Similarity 48.1%; Pred. No. 2.4e-171;
 Matches 453; Conservative 144; Mismatches 235; Indels 110; Gaps 21;
 QY 1 MLLLLGILLAFAGRTAGG--SEPEREVVPIKLDPDINGRRYYWRGP-EDSGDQGLIPQI 57
 Db 20 LLLLSITMLLCARGAHGRPTDEELVLP-SLE-----RAPHGDSITTLRL--RL 66
 QY 58 TAFQDFVILHTPDPAQFLAPAFSTHGLGVPLQGLTGGS-----SDLRRCFYSGVNA 109
 Db 67 DAFGOQLHLKLPDQSGFLAPGFTLTQV-----GRSPGSEAOHLDPDGLAHCFYSGIVNG 121
 QY 110 EPDSPAAYSLCGLGAFGYGAEEYVISPLPNAS-----APAAQRNSQGA----HLLRRG 161
 Db 122 DFGSAALSLCEGVARGVAFYLOGEEFFIQAPGVATERLAPAVPEESSARQFHILRR- 180
 QY 162 VPGSGDPTSCGVASGWNPAIRLALDPYKPRRAGFESRRR----- 206
 Db 181 -----RRSGGAKCGVMD-----DETLPSTDSRPESQNTNQMVFDPDTPQDAGK 226
 QY 207 -----SGRAKRFVSPRYVETLVVADESVMKFGADLEHYLLTLLATAARVHRPISILNPI 262
 Db 227 SGPGSIRKRRVSSPRYVETLVVADQSMADFGSLGKHYLLTLFVAAEFYKHFSIRNSI 286
 QY 263 NIVVVKVLLLRDRSGPKVTGNAALTNRNFCAMQKLNKVDKHEPYMDIAILFTRQDLC 322
 Db 287 SLVVVKILVIYEEQKGPVTSNAALTNRNFCAMQKLNKVDKHEPYMDIAILFTRQDLC 346
 QY 323 GATTCDTLGMADVGTMCDPKSCSVIEDDGLPSAFTTAHELGHVFNPHNVKVCVEVFG 382
 Db 347 GSHTCDTLGMADVGTMCDPKSCSVIEDDGLPSAFTTAHELGHVFNPHNVKVCVEVFG 406

34	307	5.9	789	1	AD07_RAT	Q63180	rattus norv
35	300	5.8	413	1	ACLA_AKAC	Q9pw35	agkistrodon
36	298	5.8	480	1	DISA_TRIGA	P15503	trimeresuru
37	289	5.6	769	1	AD11_HUMAN	O75078	homo sapien
38	288	5.6	788	1	AD07_MOUSE	O35227	mus musculus
39	286	5.5	414	1	HRTD_CROAT	P15167	crotalus at
40	280	5.4	773	1	AD11_MOUSE	Q9rlv4	mus musculus
41	277	5.4	814	1	AD15_HUMAN	Q13444	homo sapien
42	277	5.4	935	1	AD22_XENLA	O42596	xenopus lae
43	277	5.4	478	1	HRTE_CROAT	P34182	crotalus at
44	268	5.2	797	1	AD33_MOUSE	Q923w9	mus musculus
45	264	5.1	729	1	AD21_MOUSE	Q9ji76	mus musculus

OM protein - protein search, using sw model
Run on: May 9, 2003, 15:18:36 ; Search time 15 Seconds
(without alignments)
2626.834 Million cell updates/sec

Title: US-09-965-631-4
Perfect score: 5162
Sequence: 1 MLLGLITLAFAGRTAGGSE.....DOCNLHRRKPOELDFCVLRPC 950

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	2486.5	48.2	967	1	ATSL_HUMAN	Q9uh18 homo sapien
2	2485.5	48.1	968	1	ATSL_MOUSE	P97857 mus musculus
3	2474	47.9	967	1	ATSL_RAT	Q9uq79 rattus norv
4	2138.5	41.4	890	1	AT88_HUMAN	Q9uq79 homo sapien
5	2115	41.0	905	1	AT88_MOUSE	P57110 mus musculus
6	1965	38.1	1629	1	AT59_HUMAN	Q9p2n4 homo sapien
7	1915	37.1	837	1	AT54_HUMAN	O75173 homo sapien
8	1856	36.0	930	1	AT55_MOUSE	Q9r001 mus musculus
9	1847.5	35.8	930	1	AT55_HUMAN	Q9una0 homo sapien
10	1745.5	33.8	630	1	AT54_RAT	Q9esp7 rattus norv
11	1296	25.1	1077	1	AT10_HUMAN	Q9h324 homo sapien
12	1275	24.7	1593	1	AT12_HUMAN	P58397 homo sapien
13	1231	23.8	1211	1	AT52_HUMAN	O95450 h adams-2
14	1215.5	23.5	997	1	AT57_HUMAN	Q9ukp4 homo sapien
15	1192	23.1	1205	1	AT52_BOVIN	P79331 b adams-2
16	1184.5	22.9	1205	1	AT53_HUMAN	O15072 homo sapien
17	1048	20.3	860	1	AT86_HUMAN	Q9ukp5 homo sapien
18	601.5	11.7	245	1	AT54_BOVIN	Q9tt93 bos taurus
19	599	11.6	207	1	AT55_BOVIN	Q9tt92 bos taurus
20	381.5	7.4	450	1	AT10_MOUSE	P58459 mus musculus
21	378.5	7.3	824	1	AD08_HUMAN	P78325 homo sapien
22	368	7.1	956	1	AD19_HUMAN	Q9h013 homo sapien
23	362	7.0	776	1	AD38_MACFA	O9xsl5 macaca fasc
24	349.5	6.8	813	1	AD33_HUMAN	Q9bz11 homo sapien
25	343.5	6.7	571	1	DIS2_BOVJA	P30431 bothrops ja
26	342.5	6.6	774	1	AD28_MOUSE	Q9jln6 mus musculus
27	340	6.6	775	1	AD28_HUMAN	Q9ukq2 homo sapien
28	337	6.5	857	1	AD22_MOUSE	Q9rlv6 mus musculus
29	334	6.5	920	1	AD19_MOUSE	Q9p0k1 homo sapien
30	332	6.4	906	1	AD22_HUMAN	Q61824 mus musculus
31	324	6.3	903	1	AD12_MOUSE	Q05910 mus musculus
32	322	6.2	826	1	AD08_MOUSE	Q05910 mus musculus
33	314	6.1	909	1	AD12_HUMAN	O43184 homo sapien

ALIGNMENTS

RESULT 1

ATSL_HUMAN STANDARD; PRT; 967 AA.
ID ATSL_HUMAN Q9UH18; Q9UH83; Q9P2K0; Q9NSJ8;
AC Q9UH18; Q9UH83; Q9P2K0; Q9NSJ8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
..DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).
DE ADAMTS1 OR METH1 OR KIAA1346.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
RT "Cloning, characterization and mapping on human chromosome 21 of the
RT orthologue of murine Adamts-1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC TISSUE=Heart;
RX MEDLINE=99367466; PubMed=10438512;
RA Vazquez F., Hastings G., Ortega M.A., Lane T.F., Oikemus S.,
RA Lombardo M., Iruela-Arispe M.L.;
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RL family of proteins with angio-inhibitory activity.";
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells;
RX MEDLINE=20247184; PubMed=10785405;
RA Glienke J., Schmitt A.O., Pillarsky C., Hinzmann B., Weiss B.,
RA Rosenthal A., Thierauch K.H.;
RT "Differential gene expression by endothelial cells in distinct
RL angiogenic states.";
RN [4]
RP SEQUENCE FROM N.A.
RX Eur. J. Biochem. 267:2820-2830(2000).
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Park H.-S., Toyoda A., Ishii K., Taudien S., Blechschmidt K., Polley A.,
RA Ohki M., Takagi T., Sakaki Y., Lehmann R., Patterson D.,
RA Menzel U., Delabar J., Kumpf K., Zimmermann W.,
RA Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,

RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Schaefer M., Becken O., Desario A., Reichelt J., Kauter G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzaym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.,
RA "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [6]
RN PSEQUENCE OF 418-967 FROM N.A.
RC TISSUE=Melanoma;
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER (BY SIMILARITY). HAS ANGIOGENIC INHIBITOR
CC ACTIVITY. ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE.
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1938-GLU-1-LEU-1939
CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC -!- COPACITOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF170084; AAF15317.1;
DR EMBL; AF060152; AAD48080.1; ALT_INIT.
DR EMBL; AF07664; AAF3772.1;
DR EMBL; AB037767; BAA92584.1; ALT_INIT.
DR EMBL; AP001697; BAA9502.1;
DR EMBL; AL162080; CAB82413.1;
DR MEROPS; M12.222;
DR Genew; HGNC:217; ADAMTSL1.
DR MIM; 605174;
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_MTpeptidse.
DR Pfam; PR00090; tsp_1; 6.
DR Pfam; PF01421; Reprolysin; 2.
DR Pfam; PF01562; Pep_M12B_propep; 2.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 49
FT PROPEP 50 252
FT CHAIN 253 967
FT SITE 198 198
FT METAL 401 401
FT ACT_SITE 402 402
FT METAL 405 405
FT METAL 411 411
FT DOMAIN 476 559
FT DOMAIN 560 616
FT TSP TYPE-1 1.

DOMAIN	617	724	CYS-RICH.
FT DOMAIN	725	849	SPACER.
FT DOMAIN	850	908	TSP TYPE-1 2.
FT DOMAIN	909	967	TSP TYPE-1 3.
FT DOMAIN	843	846	POLY-LYS.
FT CARBOHYD	547	547	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	720	720	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	764	764	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT	227	227	P -> A (IN REF. 4 AND 5).
FT CONFLICT	468	468	O -> H (IN REF. 1).
FT CONFLICT	561	561	S -> N (IN REF. 1).
FT SEQUENCE	967 AA:	105383 MW;	CI89389324741ED1 CRC64;
Query Match	48.2%;	Score 2486.5;	DB 1; Length 967;
Best Local Similarity	49.0%;	Pred. NO. 1.2e-167;	
Matches	486;	Conservative 154;	Mismatches 250;
			Indels 101; Gaps 24;
QY	1	MLLLGLTLTAFAGRTAGSGSEPERVVRIRLPDINGRRYWRKPEDSGDQGLFIQTAF	60
DB	36	LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTRRLRHAF	81
QY	61	QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGSSDLRRCFYSGDVNAEPDSF	114
DB	82	DQQLDLRLPDSFLAPGFTLQNGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA	136
QY	115	AAVSLCGGLRGAFYGAEGYEVISPLPNAS---APAAQRNSQGA---HLLQ---ERGVPGG	165
DB	137	AALSCLGVRGAFYLLGEAYFTQPLPAASERLATAAPGKPPAPLOFHLLRNRODVG	196
QY	166	PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRRAGFSGESRRRS	207
DB	197	TCGVDDPRPTCKAETDEDETEGEDEGPQWS-----PODPALOGVGP-TGTGS	247
QY	208	GRAKRFVSIPIRYVETLIVADESMVKFAGDALEHLYTLATARLYRHPISILNPIVV	267
DB	248	IRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTFSVAARLYKHPISIRNSLVVV	307
QY	268	KVLLRRDRSGPKVTGNAALTENFCAWOKKLNKVSQKHPYWDYTAITLFTRODLCAATC	327
DB	308	KILVIHDEQKGEVTSNAALTNRFCNWQKHPPDRDAEHYDTAILFTRODLCSGTC	367
QY	328	DTLGMADVGMCDPKRSCSVIEDDGLFSAPTTAHELGHVFNPHDNYKCEEVFGKLRAN	387
DB	368	DTLGMADVGVCDPSRSCSVIEDDGLQAFTTAHELGHVFNPHDDAKQKASLNGVQDS	427
QY	388	HMSPTLIQIDRANPWSACSAIITDPLDSGHGCLLDQSPKISLPEDLPAGSYTLSCQ	447
DB	428	HMASMLSLNLDHSQPSWSPCSAYMITSLFDNGHGECLMDKPNQPIQLPGDLPSTYDANRQ	487
QY	448	CELAFGVGSKPCP-YMOYCTKLCWTKGAKGOMVCOTRHPFWADGTCGEGKCLKGACVE	506
DB	488	CQFTFGEDSKHCPDAASTCTLWCTGTSGGVLYCQTKHFPWADGTCGEGKWCINGKCVN	547
QY	507	RHNLNKH---RVDGSAKWDPPYPCSRCTCGGGVOLARRCQTNPTPANGKYCGGVKRYR	563
DB	548	KTD-RKHEDTPFHGSGWGMWGPWGDSCRTCGGVQVYTRCEDNPPVKNKGKYGKRVYR	606
QY	564	SCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCKLICRAN	623
DB	607	SCNLEDCPDN-NGKTFREOCEAHNEFSKASFGSGPAAVEWIPKYAGVSPDKCKLICQAK	665
QY	624	GTGVFYVLAPKVDGTLCSPDSTSVCGVQGCIRKAGCDGNLGSKKRFPDKCGVCGDNKSK	683
DB	666	GIGYFFVLPKVDGTPCSPDSTSVCGVQGCIRKAGCDRIIDSKKKFKDKGCGVCGNGSTCK	725
QY	684	KVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLICDDNYLALKNKSQKYLNGHFV	743
DB	726	KISGVSITSAKPYHDIITPTGATNIEVKQRNQRNGSFLAKAADGTYILNGDYTL	785
QY	744	SAVERDLVVKGSLRLRYSGTGTAVESLOASRPILPELVLSVGKMTPPRVRYSPFLPKE	803
DB	786	STLEQDLMYKGVVLRYSGSSAALRIRSFSPKLETLITQVLTVGNALRPKIKYTFYEVKK	845

QY 804 PREDKSSHPKDPGPSVLHNSVLSSNOVEQPDPRPARNVAGSKGPCSASGSGLQKRA 863
 Db 846 ---KES-----FNAITFS-----ANVIEENGECSCSGELGWORRL 878
 QY 864 VDCRGSGAGTVPACDAH--RPVETQAGE-PCPTWELSAWSPCKSCGGRGFORSLKC 920
 Db 879 VECRDINGQ---PASECAKVKFPASTRPCADHPCPQWLGEWSCSKTCGKGYKKRSLKC 935
 QY 921 VGHGRLLRDQCNLHRKPOE-LDFCVLRPC 950
 Db 936 LSHDGGVLSHESCDPLKKPKHFDFCTMAEC 966

RESULT 2
 ATSL_MOUSE STANDARD; PRT; 968 AA.
 ID ATSL_MOUSE
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
 GN ADAMTS1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RA MEDLINE=98110583; PubMed=9441751;
 RX Kuno K., Lizasa H., Ohno S., Matsushima K.;
 RT "The exon/intron organization and chromosomal mapping of the mouse
 ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";
 RL Genomics 46:466-471(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97150761; PubMed=8995297;
 RX Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
 RA Matsushima K.;
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-
 disintegrin family protein with thrombospondin motifs as an
 inflammation associated gene.";
 RL J. Biol. Chem. 272:556-562(1997).
 [3]
 RP CHARACTERIZATION, AND MUTAGENESIS OF GLD-403.
 RX MEDLINE=99303567; PubMed=10373500;
 RA Kuno K., Terashima Y., Matsushima K.;
 RT "ADAMTS-1 is an active metalloproteinase associated with the
 extracellular matrix.";
 RL J. Biol. Chem. 274:18821-18826(1999).
 [4]
 RP FUNCTION.
 RX MEDLINE=20389568; PubMed=10930576;
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
 RA Ohno H., Matsushima K.;
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";
 RL FEBS Lett. 478:241-245(2000).
 [5]
 RP FUNCTION, AND INDUCTION.
 RX MEDLINE=20243757; PubMed=10781075;
 RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
 RA Richards J.S.;
 RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
 cathepsin L proteases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
 CC "FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
 INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
 SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
 VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
 CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
 SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: CLEAVES AGGECAN AT THE 1691-GLU-|-LEU-1692

CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
 CC COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
 CC MATRIX.
 CC -!- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY
 CC INTERLEUKIN-1, OR IN VIVO IN KIDNEY AND HEART BY
 CC LIPOPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA
 CC CELLS OF PREOVULATORY FOLLICLES.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12b.
 CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
 CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 7.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AB001735; BAA24501.1; ALT INIT.
 CC DDB: D67076; BAA11088.1; ALT_FRAME.
 CC MEROPS: M12.222; -.
 CC MGD: MGI:109249; Adamts1.
 CC InterPro: IPR001762; Disintegrin.
 CC InterPro: IPR002870; Pep_M12b_propep.
 CC InterPro: IPR001590; Reprolysin.
 CC InterPro: IPR000884; TSP1.
 CC InterPro: IPR000130; Zn_Mtpeptase.
 CC Pfam: PF00090; tsp_1; 3.
 CC Pfam: PF01421; Reprolysin; 1.
 CC Pfam: PF01562; Pep_M12b_propep; 1.
 CC SMART: SM00209; TSP1; 3.
 CC PROSITE: PS00215; ADAM_MEPRO; 1.
 CC PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC PROSITE: PS00092; TSP1; 3.
 CC PROSITE: PS00142; ZINC_PROTEASE; 1.
 CC HydroLase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 CC Repeat; Extracellular matrix; Heparin-binding.
 CC SIGNAL 1 48 POTENTIAL.
 CC PROPEP 49 253
 CC CHAIN 254 968
 CC SITE 206 206
 CC METAL 402 402
 CC ACT_SITE 403 403
 CC METAL 406 406
 CC METAL 412 412
 CC DOMAIN 477 559
 CC DOMAIN 560 617
 CC DOMAIN 618 725
 CC DOMAIN 726 850
 CC DOMAIN 851 909
 CC DOMAIN 910 968
 CC DOMAIN 910 968
 CC DOMAIN 195 199
 CC CARBOHYD 548 548
 CC CARBOHYD 721 721
 CC CARBOHYD 765 765
 CC CARBOHYD 783 783
 CC CARBOHYD 946 946
 CC MUTAGEN 403 403
 CC CONFLICT 335 335
 CC CONFLICT 425 425
 CC SEQUENCE 968 AA; 105841 MW; 42EBDA55499FB6C1 CRC64;
 Query Match 48.1%; Score 2485.5; DB 1; Length 968;
 Best Local Similarity 48.0%; Pred. No. 1.4e-167;
 Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

```
QY 1 MLLGLTLAFAGTAGG--SEPEREVVVVIRLDPDINGPRYYWFGP-EDSGDQGLIFQI 57
Db 37 LLLLASITMLCARAGAPTEDEELVLP-SLE-----RAPHGDSITTLRL--RL 83
QY 58 TAFQEDFYHLTPDAQFOLAPAFSTHGLGVPLQGLTGGSS-----SDLRKCFFYSGDVNA 109
Db 84 DAFGQQLHLKLQPDGSLFAPGTLQTV-----GRSPGSEAOHLDPGDLAHCFYSGTVNG 138
QY 110 EPDSFAVNSLGGLRGARGFYGAEEVYISPLPNAS-----APAAQNSQGA-----HLLQBRG 161
Db 139 DFGSAALSLCEGVGAFYLOGEFFIOPAPGVATERLAPAVEEESARPOFHILRRR- 197
QY 162 VPGPGSGDPTSCGVASGNPAILRALDPYKPRRAGFESRRR-----DETLPDSRPESQNRNQWVPDRPTPDACKP 243
QY 207 -----SGRAKRFVSIPIRYETLVVADESVMKFGADLEHYLLTLATARLYRHPSILNPI 262
Db 244 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSLKHVLLTLFVSAAFYKHPSIRNSI 303
QY 263 NIWVKVILLRDBDQPKVTGNAALTIRNFCWOKKLVSDKHBEYNDTALTITRQDLC 322
Db 304 SLVVKILYIEBQKQPEVTSNAAITLRNFCWQKHNSPDSRDEHYDTALTITRQDLC 363
QY 323 GATTCTGLMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNNVKVCEYFG 382
Db 364 GSHTCTGLMADVTGCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDNAKHCASLNG 423
QY 383 KLRANHMGSPTLIQIDRANPWASCSAAITDFLDGSHGDCILDOFSKPLSPEDLPQASY 442
Db 424 VTGDSHLMASMLSSLDHSPQSPCSAYMVTSLFDNGHGECLMDKPNPKLPSPDLPGTLY 483
QY 443 TLSQCELAGFVGSKPCP--YMOYCTKLWCTGKAGOMVCTRHFWADGTSGEGKLCIK 501
Db 484 DANQCOFTGEESKHCPCDAASTCTLWCTGTSGLLVCQTKHFPWADGTSGEGKWCVS 543
QY 502 GACVERHNLNKH---RVDSWAKWDYPGPCSTCTCGGGVGLARQCTNPTTPANGKYCEGV 558
Db 544 GKCVNKTDM-KHFAATPVHGSWGPWGPCSRCTCGGGVQYVTRCEDNVPVKNKGKYCEGK 602
QY 559 RVYRSCNLEPCSPSSASGSKSFREOCEAENGYNHSTNRLTLAVAKVPKYSVSPDKCKL 618
Db 603 RVYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKSGFNEPTVEWTPKYAGVSPDKCKL 661
QY 619 ICRANGTYPIYVLAPKVDGTLGSPDSTSVYQGGKIKAGCDNLSGKKRFPKCGVCGGD 678
Db 662 TCEAKGIGYFVLQPKVDGTPCSPDSTSVYQGGVQCVKAGCDRIIDSKKKFDKCGVCGGN 721
QY 679 NKCKKVTGLTTPMHGYNFVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQGYLLN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVIPAGATNIEVKHNRQSRNNGSFLAIRAADGTIYLN 781
QY 739 GHFVSAVERDLVVKSLRYSGTGTAVESLQASRPILPLETLVEVLSVGKMTPPRVRYSF 798
Db 782 GNFTLSLEODLYTKGVLYRSGSRALEIRSFSPKLEPTLIQVLMVGHALRPKTKFY 841
QY 799 YLKPREDKSSHPPKDRPGPSVLHNSVLSLNOVEQPDPRPARPVAWGSWPCSCSGG 858
Db 842 FMKKTES-----FNAIPFS-----EWVTEEWGECSTKCGSG 874
QY 859 LQRAVDCRSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSGRGFOR 915
Db 875 WQRVVQCRDINGH---PASECAKEVKPASTRCPADLPCPHWQVGDWSPCSKTCGKGYKK 931
QY 916 RSLKCVGHGRLLARDOCNLRHXPQE-LDFCVLRPC 950
Db 932 RTLKCVSHDGVLSNESCDLPKPKKHWFIDFCTLTQC 967
RESULT 3
ATSL_RAT
ID_ATSL_RAT STANDARD; PRT; 967 AA.
AC Q9WUQ1; Q9ER11;
```

```
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN ADAMTS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Brain;
RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA Little S.P.;
RT "Induction of a disintegrin and metalloproteinase with the
RT thrombospondin type I motif (ADAMTS).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE OF 18-967 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luehthi M., Hoesli M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT endothelial cells in cirrhotic rats.";
RL Liver 20:165-172(2000).
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1683-GLU-1-LEU-1684
CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -!- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM
CC CIRRHOTIC LIVER.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF149118; AAD34012.1; -.
EMBL; AF304446; AG29823.1; -.
MEROPS; M12.222; -.
InterPro; IPR001762; Disintegrin.
InterPro; IPR002870; Pep_M12B_propep.
InterPro; IPR001590; Repolysin.
InterPro; IPR000884; TSP1.
InterPro; IPR000130; Zn_Mtpeptdse.
Pfam; PF00090; Tsp1; 6.
Pfam; PF01421; Repolysin; 2.
Pfam; PF01562; Pep_M12B_propep; 2.
SMART; SM00209; TSP1; 3.
PROSITE; PS0215; ADAM_MEPRO; 1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
PROSITE; PS00092; TSP1; 2.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
```

FT SIGNAL 1 54
 FT PROPEP 55 252
 FT CHAIN 253 967
 FT SITE 205 205
 FT METAL 401 401
 FT ACT_SITE 402 402
 FT METAL 405 405
 FT METAL 411 411
 FT DOMAIN 476 558
 FT DOMAIN 559 615
 FT DOMAIN 616 724
 FT DOMAIN 725 857
 FT DOMAIN 858 907
 FT DOMAIN 908 967
 FT DOMAIN 194 198
 FT CARBOHYD 547 547
 FT CARBOHYD 720 720
 FT CARBOHYD 764 764
 FT CARBOHYD 782 782
 FT CARBOHYD 945 945
 FT CONFLICT 21 21
 FT CONFLICT 26 31
 FT CONFLICT 49 49
 FT CONFLICT 72 72
 FT CONFLICT 79 79
 FT CONFLICT 249 249
 FT CONFLICT 262 265
 FT CONFLICT 607 607
 FT CONFLICT 936 936
 FT CONFLICT 962 962
 FT CONFLICT 967 967
 FT SEQUENCE 967 AA; 105705 MW; F93C864F6DCDB4CF CRC64;

Query Match 47.9%; Score 2474; DB 1; Length 967;
 Best Local Similarity 48.4%; Pred. No. 9,2e+167;
 Matches 477; Conservative 162; Mismatches 255; Indels 92; Gaps 22;

QY 1 MLLGILLTAPAGTAGG--SEPEEVVVPRLDPDINGRRYYWRGPEDSQDGLIFQIT 58
 DB 37 LLLLASITMLLCVGAHGRTEDEELV-----PSLERARGH-----DSTILLRLD 83
 QY 59 AFOEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGG--SRLRRCFYSSDVNAE 110
 DB 84 AFGQQLHLKLPQDSFLAPGTLQTV-----GRSPGEAHLDPGLAHCFYSGTVNGD 138
 QY 111 PDSFAVSLCGGLRGAFGYGAEYVISPPLNAS-----APAAQRNSQA-----HLL--QRR 160
 DB 139 PSSAALSILCEVGRGAFYVQLQEEFPIQAPAVATERLWPAEPKEESJAPPRFHILRRRR 198
 QY 161 GVPGFSGD-----PTSRCGVASGNPAILRALDPKPRAGFGRSRRRRGRKRFV 214
 DB 199 GSGGAKCGVMBDETLPTSSNGRESQNTPDQWPLRNP--TPQAG--KPTGPGSIRKRFV 254
 QY 215 SIPRVEVTLVADESVMKFGADLEHLLTLLATAARLYRHPSTILNINIVVKKVILLRD 274
 DB 255 SSPRIVETMLVADQSMADFHGSLGKHLTLTFLSVAARFYKHPSTIRNSLSLVVKKILYIE 314
 QY 275 RDSGPKVTGNALTLRNFCAWKLNKVSDEKHPYWDTAIFLTRQDLICGATTCTDLGNAD 334
 DB 315 EQKGPEVTSNAALTILRNFCSQKQHNPSDRDPEHYDTAIFLTRQDLICGSHCTDLGNAD 374
 QY 335 VGTMCDPKRRSSVIEDDGLPSAFTAHGLHGVFNPNHNDNVKVCVEVFGKLRANHMSPFL 394
 DB 375 VGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHNDHDAKAGFNGVSGDSHLMASML 434
 QY 395 IQIDRANPWSACSAIITDFLDSHGDCLLDQPSKPTSLPDLPGA-YTLSQCELAGV 454
 DB 435 SSLDSHQSPWSPCSATMYVTSFDNNGHECLMDKPNKPLPSDLPTGLYDANRQQQFTGE 494
 QY 455 GSKPCP-YMQVCTKLVCTGKAGQWVCOTRHFHPWADGTSCEGSKLCLKAGACVERHNLKH 513
 DB 495 ESTHCPDRASTCLWCTGTSGLLVCOQTKHFPWADGTSCEGKWCYSGKCVNKIDM-KH 553
 QY 514 ---RVDGSWAKWDPPGCSRTCGGSQLARRQCTNPTPANGGKYCEGVRYKRYSCNLEPC 570

DB 554 FATPVHSGWGPWGPWGDSCRTCGGVQYTMRECDNPVKNRGKYCEGKRYRYRSCNIEDC 613
 QY 571 PSSASGKSPREOCEAFNGVYNHSTNRLTLAVANVPKYSVGSVPDRCKLCLCRANGTCYFYV 630
 DB 614 PDN-NGKTFREOCEAHNEFKASFGNEPTVETWPKYAGVSPDRCKLCLCEAKGIGYFFV 672
 QY 631 LAPKVVDDGTLCSPDSTSVCVQGVQKICRAGCDNGLSGKRRFDKCGVCGDNKSKKVIULET 690
 DB 673 LQPKVVDDGTLCSPDSTSVCVQGVQKICRAGCDNGLSGKRRFDKCGVCGDNKSKKVIULET 732
 QY 691 KPMHGNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNQSKYLLNHHFVVSAYVREL 750
 DB 733 STRPGYHDIVTIPAGATNIEVKHNPGRSNGSFLAIRAADGTYILNGFTLSTLEQDL 792
 QY 751 VVKGSLRLYSGTGTAVESLQASRPILPLETVLSVKGMTPPRVRYSFYLPK--EPREDK 808
 DB 793 TYKGTVLRYSGSSAALERIRSFLEPLTIQVLMVGHALRPKIKTYFMKKTEH----- 848
 QY 809 SSHPKDPRGPSVLHNSVLSLNQVEQPDPRPARVAGSNWGPCSASCGLQKRAYDCRG 868
 DB 849 -----FNAIPTFS-----EWIEEWGECSTKCGSQWQRRVVECRD 883
 QY 869 SAGQIVPACDAH--RPVETQACGE--PQVWELSAWSPCKSCGCGRFGRRSLKCVGHGG 925
 DB 884 INGH---PASECAKEVKPASTRPCADLPCPRWQGVGDMSPCKTCGKGYKRTLKCLSHDG 940
 QY 926 RLLARQCNLHRKPOE--LDPCVLRPC 950
 DB 941 GVLNSCSDPLKPKKHYDFCILTQC 966

RESULT 4
 AT58_HUMAN
 ID AT58_HUMAN STANDARD; PRT; 890 AA.
 AC Q9UP79; Q9NZS0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2) (METH-8).
 DE Homo sapiens (Human).
 GN ADAMTS8 OR METH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=99367466; PubMed=10438512;
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S., Lombardo M., Iruela-Arispe M.L.;
 RA "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity";
 RL J. Biol. Chem. 274:23349-23357(1999).
 RN [2]
 RP SEQUENCE OF 195-440 FROM N.A.
 RX MEDLINE=20079168; PubMed=10610729;
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
 RA "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";
 RL Genomics 62:312-315(1999).
 CC -!- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES.
 CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADULT AND FETAL LUNG, LOWER EXPRESSION IN BRAIN, PLACENTA, HEART, STOMACH AND FETAL BRAIN AND KIDNEY.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY

Genomics 62:312-315(1999).

-1- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES (BY SIMILARITY).

-1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).

-1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN ADULT LUNG AND HEART AND LOW EXPRESSION DURING MOUSE DEVELOPMENT.

-1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.

-1- PTM: THE PRECURSOR IS CLEAVED BY A FUIN ENDOPEPTIDASE (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.

-1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.

-1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF175282; AAF25805.1; .

HSSP; P34179; I1AG.

MEROPS; M12.226; .

MGD; MGI:1353468; Adamts8.

InterPro; IPR001762; Disintegrin.

InterPro; IPR001590; Reprolysin.

InterPro; IPR000884; TSP1.

InterPro; IPR000130; zn_MTPeptdse.

Pfam; PF00090; tsp.1; 2.

Pfam; PF01421; Reprolysin; 1.

SMART; SM00209; TSP1_2.

PROSITE; PS00142; ZINC_PROTEASE; 1.

PROSITE; PS00215; ADAM_MPRO; 1.

PROSITE; PS50092; TSP1; 1.

PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.

KW Hydrolyase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;

KW Repeat; Extracellular matrix; Heparin-binding.

FT SIGNAL 1 28

FT PROPEP 29 228 BY SIMILARITY.

FT CHAIN 229 905 ADAMTS-8.

FT METAL 378 378 ZINC (CATALYTIC) (BY SIMILARITY).

FT ACT_SITE 379 379 BY SIMILARITY.

FT METAL 382 382 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).

FT DOMAIN 453 541 DISINTEGRIN-LIKE3.

FT DOMAIN 542 598 TSP TYPE-1 1.

FT DOMAIN 599 705 CYS-RICH.

FT DOMAIN 706 847 SPACER.

FT DOMAIN 848 905 TSP TYPE-1 2.

FT CARBOHYD 415 415 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 480 480 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 506 506 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 615 615 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 905 AA; 98879 MW; 124D4132B33A0CAE CRC64;

Query Match 41.0%; Score 2115; DB 1; Length 905;

Best Local Similarity 45.6%; Pred. No. 1.8e-141;

Matches 431; Conservative 150; Mismatches 259; Indels 106; Gaps 25;

QY 2 LLLIGLTL----AFAGRTAG---GSPEREVVPVRLDPDINGRRVYWRGPDSDGQGLI 54

DB 13 LLLLLLQPPPLVCGAPAGTGAQAS-ELVVPTL-----PGSASE--LA 56

QY 55 FQTAFQEDFVHLTPDPAFLAPFSTEHLGVPLOGLTGSSDLRRCFYSGDVNAEPDSF 114

DB 57 FHLSPAGQGVFLRLAPDASFLAPEFKIERLG-GSSAAGGEPGLRGCFSGTNGERESL 115

QY 115 AVSLCGGLRGAFGYGAYVIVSPILPNASAPAAQNSOGA-----HLLQRGVPGGFS 167

DB 116 AAMSCVAGSGSFLGAGEFTIQP-----QCAGDSLDPQHLRWG-PGQRR 161

QY 168 GDTPTSCGVASGNPAILLRALDYPKPRAGGESRS-----RRSG----- 208

DB 162 EDP-----GLAAEVFPLPQGLEWEEMGNGQQRSDNEEDRKQDKEGLLKETEDSRKVP 217

QY 209 -----BAKRVSITPRYVETLVVADESVMKFGADLEHYLLTLATARLYRHPISILNP 261

DB 218 PFGSKTSKRVSARFVEILLVADASMAAFYGLQNHILITVMSMAARIYKHPISIRNS 277

QY 262 INIVVVKLLLRDRDSGPKVTGNAALTNRFCAMOKKLNKYSDKHPEYWDTAIFTTODL 321

DB 278 VNLVVVKVLIVKRWGPEVSDNGGLTLRNFCSWQRRENKPSDRHPEHYDTAILFTRQNE 337

QY 322 CG-ATTCDTLGMADVTGMDPKRSCSVTEDDGLPSAFTTAHELGHVFNPHNDNVKVCVEV 380

DB 338 CGKGQCQDTLGMADVTGMDPKSCSVIKDEGLQAAYLAHELGHVLSMPHDDSKPCVRL 397

QY 381 FGKLRANHMSPTLQIDRANPWSACSAITTFDLDGSHGDCLLDOPSKPISLSPDLPGA 440

DB 398 FGPMPKYHMAFFFIHVAKTLPWSPCSAVYITLDDGSHGDCLLDAPTSVLPPTGLPGH 457

QY 441 S--YTLSSOCELAFGVGSKPCP---YMOYCTKLMTGKAKGMVQCOTRH---FFWADGISC 493

DB 458 STLVELDQCKQIFGPDFRHCNPNTSVEDICVQLCARHRDSDPEICHTKNGSLLLWADGIPC 517

QY 494 GEGKCLLAGACVERHNLN--KRVDSGSAKWDPYGPCSRTCGGVQLARROCTNVPFANG 551

DB 518 GFGHCLDGSCLVEDVENPKAVVDGMPWPMGQCSTCGGGTQFSNRECDNPNPONG 577

QY 552 GKYGCVRVKYRSCNLEPCSSAGKSPREOCFAFNYNHSTNRLTLAVAVMKYKSGVS 611

DB 578 GFGCLGERVKYQSCNTECP--PNGKSPREOCCEKYNAINH--TLDGNTFLQWPKYKSGVS 634

QY 612 PRDCKLICRANGTGYFYVLAPVVDGTLCSPDSTSVCGVKCIKAGCDGSLGSKKRFDK 671

DB 635 PRDCKLFCRARGSEFKVEAKVIDTGLCGPDTLSICVRGQCVKAGCDHVNVSPKLDK 694

QY 672 CGVCGGDNKSKVKTGLTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNDYLAIKNS 731

DB 695 CGVCGGKTACKIKSGSFTPEFYGYNDIVTIPAGATNIDVKQRSHPGVRNDGSLAUKTA 754

QY 732 OGKYLNGHFVYSAVERDLVVKSLRYSGTGTAVESLQASRPILPTLTVELSV-CKMT 790

DB 755 NGVYLLNGLAISALEQDILVKGILKYSGNATLERLQSFQALPEPLTVQLLTVSVEVF 814

QY 791 PPRVRYSTLPKPEPREDKSSHPKDPGRGFSVLHNSVLSLNSQVEQDPPRPPARWVAGSWG 850

DB 815 PKVRYTFEVD--MDFSQNSKERATT---NIIQSL-----PSAEWVLGDWSE 859

QY 851 CSASGSLQKRAVDCRGSAGQRTVPACDAHRPVEYQACG-EPCP 895

DB 860 CPSTCRGSQWRRTVECDPDSQOAS-DTCDEALKPEDAKPCGSPCP 904

RESULT 6

ATS9_HUMAN

ID ATS9_HUMAN STANDARD; PRT; 1629 AA.

AC Q9P2N4; Q9N29;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).

GN ADAMTS9 OR KIAA1312.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (SHORT ISOFORM).

RC TISSUE=Fetal;

RX MEDLINE=20396138; PubMed=10936055;

RA Clark M.E., Kelner G.S., Turbeville L.A., Boyer A., Arden K.A.,

RA Maki R.A.;
 RT "ADAMTS 9, a novel member of the ADAM-TS/Metallospandin gene
 family";
 RL Genomics 67:343-350(2000).
 RN [2]
 RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
 RT Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro.*;
 RL DNA Res. 7:65-73(2000).
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.
 CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,
 CC PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN
 CC COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR
 CC THYMUS.
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
 CC SIMILARITY).
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 DR EMBL: AF261918; AAF89106.1; -;
 DR EMBL: AB037733; BAA92550.1; -;
 DR HSPB: P15167.1ATL.
 DR MEROPS: M12.021; -;
 DR MIM: 605421; -;
 DR Genew: HGNC:13202; ADAMTS9.
 DR InterPro: IPR001762; Disintegrin.
 DR InterPro: IPR002870; Pep_M12B_propep.
 DR InterPro: IPR001590; Repolysin.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR000130; Zn_MTpeptdse.
 DR Pfam: PF00090; tsp_1; 11.
 DR Pfam: PF01421; Repolysin; 1.
 DR Pfam: PF01562; Pep_M12B_propep; 1.
 DR SMART: SM00209; TSP1; 12.
 DR PROSITE: PS00215; ADAM_MEPRO; 1.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS00092; TSP1; 9.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Alternative splicing.
 FT SIGNAL 1 18
 FT PROPEP 19 287
 FT CHAIN 288 1629
 FT DOMAIN 509 587
 FT DOMAIN 589 642
 FT DOMAIN 645 752
 FT DOMAIN 753 880
 FT DOMAIN 999 1053
 FT DOMAIN 1056 1108
 FT DOMAIN 1111 1136
 FT DOMAIN 1184 1239
 FT DOMAIN 1240 1295
 TSP TYPE-1 2.
 TSP TYPE-1 3.
 TSP TYPE-1 4.
 TSP TYPE-1 5.
 TSP TYPE-1 6.

FT	DOMAIN	1332	1383	TSP TYPE-1 7.
FT	DOMAIN	1386	1439	TSP TYPE-1 8.
FT	DOMAIN	1445	1498	TSP TYPE-1 9.
FT	DOMAIN	1501	1554	TSP TYPE-1 10.
FT	DOMAIN	1562	1612	TSP TYPE-1 11.
FT	DOMAIN	88	96	POLY-SER.
FT	SITE	223	223	CYS-TER. SWITCH (POTENTIAL).
FT	METAL	434	434	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	435	435	BY SIMILARITY.
FT	METAL	438	438	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	444	444	ZINC (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	271	271	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	749	749	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	840	840	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1213	1213	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1267	1267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	1084	1072	CLVTCGKSH -> VRWEGCYFP (IN SHORT ISOFORM).
FT	VARSPLIC	1073	1629	MISSING (IN SHORT ISOFORM).
FT	CONFLICT	367	367	F -> L (IN REF. 1).
FT	SEQUENCE	1629 AA;	182649 MW;	C1C4CEFF58B941F CRC64;
Query Match		38.1%	Score 1965;	DB 1; Length 1629;
Best Local Similarity		39.9%	Pred. No. 1.4e-130;	
Matches 411;		Conservative 145;	Mismatches 351;	Indels 122; Gaps 21;
QY	22	ERVVVPIRLDP-----	DINGRRVYV-----	RGPDSDGDOGLFOITAFQ 61
DB	47	EYEVSPIRVNALGEPTNVHFKTRRSIN	ATDPWAFASSSSSSTSSQAHYRLSAFG 106	
QY	62	EDFYLHLPDAQFLAPAFSTHGLVGP----	LOGLTGGSSDLRRCFYSGDVNAEPDSFRAV 117	
DB	107	QQFLENLTANAGFIAPLFTVLLGT	PGVNTQFYSEEAELKCFYGYVNTNSHTAVI 166	
QY	118	SLCGGLRGAFYRGABYVLSPLPNASAPAA	QNSQAHLLQRRVGGPGSGDFTSRGVA 177	
DB	167	SLCGSMGLTFRSHDGYFTIEPQSMDEQ	DEEENKPHIYRRSAPQREPSTGRACOTS 226	
QY	178	SGWNPAILRALDPYKPRRAGFGE-----	SRSRRSRG 208	
DB	227	EHNK---RHSDKKTKRARKWGERIN	LAGDVAAALNSGLATEAFSAVGNKTDNTRK 283	
QY	209	RARRFVSIPRYVETLVVADESVMVKG	ADLEHYLLTLLATAARLYRHPHSILNFINVVK 268	
DB	284	RTKRFLSYPRFEVLVADNRVSYG	ENLQHYLTMSIVASIVYKOPSGIGNLINIVVN 343	
QY	269	VLLLRDRDSGPKVTGNAALT	LRNFCWOKKLNKVSDEKPEYWDALTALTRQDLCCA-TTC 327	
DB	344	LIVIHNEQDQPSISFNAQTTLKNFCQ	QWHSKNSPGGIH---HDTAVLLTRQDLCRAHDKC 400	
QY	328	DTLGMADVGTMCDPKRSVIEDDGLPS	AFATTAHELGHVFNPHDNVKNVYCEEVEFGKLRAN 387	
DB	401	DTLGLAELGTICDPYRSCSIS	EDSGLSGLSTAFITAHLEGHVFNPHDNNKCKEE-GVKSPQ 459	
QY	388	HMSPTLIQIDRANPWSACSAIIIT	DFLDSGDCDLDOP-SKPTSLPDLPGASTLSQ 446	
DB	460	HVMAPTFLNFTNPMWMSKCSRK	YITFEFLDTGEGCLNPEPSRPYPLVOLPGILYNVVK 519	
QY	447	QCELAFCVSGKPCPYMOYCTKLWC---	TGKAKGOMVQCTRHPPWADCTSCGEGKCLCKGA 503	
DB	520	QCELIFPGSQVCPYMMQCRRLN	CNWNVNVGHKG---CRTQHTFWADCTCECPGKCKYGF 576	
QY	504	CVERHNLNHRVDGSKAWKWD	PPYGPCSRTCGGVQOLARRQCTNPTPANGKCYCEGVAKYR 563	
DB	577	CVPK-EMDVPTDGSWGSWSP	FGTCSTRTCGGGIKTAIRECNRPKNGKCYCVGRMRKF 635	
QY	564	SCNLEPCPSASGKSPREOCE	AFNFGYNHNSTNRLTLAVAVPKYSGVSPDKCKLICRAN 623	
DB	636	SCNTEPCLKQ--RDFRDEQ	CAHFDGKHFNNGLLNPVRWPKYSGILMKRCKLFCRVA 693	
QY	624	GTGYFYVLAPVVDGTLCS	PDSTSVCVQKCIKAGCDGNLGSKKRDKCGVCGGDNKCK 683	

```

Db 694 GNTAYQLRDRVIDGTGCGODTNDICVQGLCRQAGCDHVLNSKARRDKCVCVGDNDSSCK 753
QY 684 KVTGLFKPMHGVNFVAIPAGASSIDIRQGVYKGLIGDNNYALNKSQCKYLLNGHFVY 743
Db 754 TVAGTFNTVHGVNTVVRIPAGATNIDVROHSFSGTDDNNYLALSSKGEFFLNGNFVY 813
QY 744 SAVERDLVVKSLRYSGTGTAVESLQASRPILFVLSVCKMTTPPRVRYSFYLPKE 803
Db 814 TMAKREIRGNNAVVEYSGSETAVERINSTRIEQELLQVLSVCKLYNFDVRYSFNPIE 873
QY 804 PREDK---SSHPKDRGP-----SVLHNSVLSLNOVEQDPRPP--- 840
Db 874 DKPOQFYWNH-----GPMQACSKPCQGERKRLVCTRESQDLTVS---DQRCDRLPQPC 925
QY 841 -----ARWAGSWGPCASGSGLOKRAVDCR-----GSAGQETVPACDAHR 883
Db 926 HITEPCGTCDLHHVVASSECSAQCGGLGRTLDIYCAKYSRLDGKTEKVDGFCSSHPK 985
QY 884 PVTQACGECPT--WELSAWSPSCSGRGFORSLKGVGHGRLLARDQCNLHRKPOE 941
Db 986 PSNRKCSGBCNTGGWRYSAWTECSKSCDGGTORRRAICVNRNDVLDKCT-HQEKVT 1044
QY 942 LDFCVLRPC 950
Db 1045 IQRCSEFP 1053

RESULT 7
AT54_HUMAN
ID AT54_HUMAN STANDARD; PRT; 837 AA.
AC Q75173; Q9UN83;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-4 precursor (PC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)
DE (ADMP-1).
GN ADAMTS4 OR KIA00688.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suvama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99286303; PubMed=10356395;
RA Tortorella M.D., Burn T.C., Pratta M.A., Abbaszade I., Hollis J.M.,
RA Liu R.-Q., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R.,
RA Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,
RA Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K.,
RA Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,
RA Tzaskos J.M., Arner E.C.;
RT "Purification and cloning of aggrecanase-1: a member of the ADAMTS
RT family of proteins.";
RL Science 284:1664-1666(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX Sawai Y., Nagase H., Saklatvala J., Clark A.R.;
RT "ADAMTS-4 genomic locus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=20400518; PubMed=10827174;

```

```

RA Tortorella M.D., Pratta M.A., Liu R.-Q., Abbaszade I., Ross H.,
RA Burn T.C., Arner E.C.;
RT "The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for
RT aggrecan substrate recognition and cleavage.";
RL J. Biol. Chem. 275:25791-25797(2000).
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. COULD ALSO BE A ONE
CC CRITICAL FACTOR IN THE EXACERBATION OF NEURODEGENERATION IN
CC ALZHEIMER'S DISEASE.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
CC site.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG AND HEART. EXPRESSED
CC AT VERY LOW LEVEL IN PLACENTA AND SKELETAL MUSCLES.
CC -!- INDUCTION: BY INTERLEUKIN-1.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FUIN ENDOPEPTIDASE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO AS ADAMTS2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: AB014588; BA313663.1;
EMBL: AF148213; AAD41494.1;
EMBL: AY04847; AAL02262.1;
MEROPS: M12.221;
GENE: HGNC:220; ADAMTS4.
MIM: 603876;
InterPro: IPR001762; Disintegrin.
InterPro: IPR001590; Reprolysin.
InterPro: IPR000884; TSPI.
InterPro: IPR000130; Zn_MTPeptidse.
Pfam: PF00090; tsp.1; 1.
SMART: SM00209; TSPI; 1.
PROSITE: PS00142; ZINC_PROTEASE; 1.
PROSITE: PS0215; ADAM_MEPRO; 1.
PROSITE: PS00092; TSPI; 1.
PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Extracellular matrix;
FT SIGNAL 1 51 POTENTIAL.
FT PROPEP 52 212
FT CHAIN 213 837
FT SITE 194 194 ADAMTS-4.
FT METAL 361 361 CYSTEINE SWITCH (POTENTIAL).
FT ACT_SITE 362 362 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 365 365 BY SIMILARITY.
FT METAL 371 371 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 437 519 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 520 576 DISINTEGRIN-LIKE.
FT DOMAIN 577 685 TSP TYPE-1.
FT DOMAIN 686 837 CYS-RICH.
FT DOMAIN 247 252 SPACER.
FT CARBOHYD 68 68 POLY-ALA.
FT CONFLICT 77 77 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 626 626 A -> T (IN REF. 1).
FT CONFLICT 682 682 R -> Q (IN REF. 3).
FT CONFLICT 682 682 G -> R (IN REF. 3).
SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;

```

Query Match

37.1%; Score 1915; DB 1; Length 837;

```
Best Local Similarity 46.3%; Pred. No. 2.1e-127;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

QY 1 MLLGILTLAFAGRTAGSGSEPERVVPIRLDPDINGRRYYWRGPEUSGGGLFIQTAF 60
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 37 LLLLLLLSLPSARLASLPREEIVPEKLNGSVL-----PGSGAPARLLCRLOAF 88
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 61 QDFVHLHTLPAOFLAPAFSTHGLVPLQGLTGSSDLRRCFYGSDVNAEPDSPAAYSLC 120
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 89 GETLLELEQDSGVQVEGLTVQYLGQAPL-LGGAEP--GYILGTINGDPESVASLHWD 145
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 121 GG-LRGAFYGYAEYVISPPLNASAPAAQRNSQGAHLLOREGVPGSGDPTSCGVASG 179
   || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 146 GCALLGLVQYRGAEHLQPLEGGTPNSA--GGPGAHILRRK-SSPAGSGQPMCNV--- 196
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 180 WNPAILRALDPKPRACFGSGRSRRSRGRKRFVSPRYVETLVVADESVMKFGHADLE 239
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 197 -----KAPLGSFSPRPR--RAKREASLSRFVETLVVADKRWAAEHGAGLK 239
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 240 HYLLTLTLATAARLYRHPSTLNPINIVYKVLRLDRDGSKPVGTGNAALTLRNFCAWOKKL 299
   |||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 240 RYLLVMAAAKAFKPSIRNPVSLVTLVILGSGEESQVGPSPSAQTLSFCANQGL 299
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 300 NKVSDKHPYWDYTLFTFRQDLGATCTDTLGMADVGMTCMDPKRSCSVIEDDGLPSAFTT 359
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 300 NTPEDSDPHFDYTLFTFRQDLGATCTDTLGMADVGMTCMDPKRSCSVIEDDGLQSAFTA 359
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 360 AHELGHVFNMPHDNVKVEEVEGKL-RANHMSPTLIOIDRANPWSACSAALITDFLDSG 418
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 360 AHELGHVFNMLHDNSKPCISLGRVPLSTSRHVNAPVMAHVDPPEWSPCSARFITDFLDNG 419
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 419 HGCDLDDQSKPISLPEDLPKASYTLTSCQELAFGVGSKPCPTM-QYCTKLWCTGKARGQ 477
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 420 YGHCLLDKPEAPLHLPTVFPFGKDYDADQCQLTFGPDNRHCPQLPPPCAAWLKCSGHLNGH 479
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 478 MVCQTHFPWADGTCGGEKGLCKGACVERHNLKRVHD--GSWAKWDPYGPCSRTCGGG 535
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 480 AMCQTKHSPWADGTPCGPAQAACMGRCRLHMDLODFENIPOAG-WGPGWPWGDSCSRTCGG 539
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 536 VOLARQCCTNPYPANGKYCEGVRYKYSCLNEPCPSASGSKPFREOCEAFNGYHNHTN 595
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 540 VQFSRDCRTPVRNGKCYCEGRFRFRSCNTEDCP-TGSALTFREOCAA---YNNHRTD 595
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 596 RLTL---AVAWPKYSGVSPDKCLICRANGTYFYVYLAQKVDGTLTCSPDSTSVQVG 652
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 596 LFKSFGPMDMVPRTYGVAPDQCKLTQARALGYVYVLEPRVVDGTPCSPDSSVCVQVG 655
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 653 KTKACGDNGLSKKRFKCGVCGGDNKSKCKVTGLTKPMHGYNFVVAIPAGASSIDIR 712
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 656 RCIHAGCDRLIGSKKFKDKCMVCGGDSGSKQSGSKFRKFRYGNVVTIPAGATHILVR 715
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 713 QRCYKGLIGDDNYLALKNSOGKYLNGHFVVSVERDLVVKGLS-LRYSGTGTAVESLQA 771
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 716 QQGNPG--HRSYIYALKLPDGSYALNGEYTLMPSTDVVLPAVSLRSGATASETLSG 773
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 772 SRPILPLAVEVLSVGKMTPTPRVRYSYFLPKPEPKEDKSSHPKD 814
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 774 HGPLAQPLTLQVLVAGNPQDTRLRYSEFFVPRPTPTSPRPTPD 816
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 8
ATSS_MOUSE
ID ATSS_MOUSE STANDARD; PRT; 930 AA.
AC Q9R001;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS) (Aggrecanase-2)
DE (ADMP-2) (Implantin).
DE ADAMTS5.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
FT CARBOHYD 807 807 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 930 AA; 101780 MW; 84DB84B26170DADC CRC64;

Query Match
Best Local Similarity 36.0%; Score 1856; DB 1; Length 930;
Matches 370; Conservative 115; Mismatches 265; Indels 106; Gaps 13;

QY 90 GLTGSSDLRCFYGDYNAEPDSPAAYSLGGLRGAFYGAEVLSPLNAPAAQR 149
Db 121 GLSASSGHRGHCYFRTGDSRSLAVFDLCGLDGFVAVKHARYTLKPLRGSWAEYER 180
QY 150 -----NSOGAHLQRG-----VPGPSPGDPTRSCGVASGWNPAILRALDPY 191
Db 181 IYDGGSSRLHVYNGEFSFEALPPRASCETPASPSPG----- 217
QY 192 KPRRAGFGESESRSSRG-----RAKRFYSIPRYVETLVVAD 227
Db 218 -POESPSVHSRRSALAPOLLDSAPSPGNAGPQWNRHRRSRASRAQVLELLVAD 276
QY 228 ESMVKFHGADLEHYLLTLLAARLYRHPSTILNPINIVVKVLLLRDSDGPKVTGNAAL 287
Db 277 SSMARYGRGLQHYLLTLLASTANLYSHASHTENHRLAVKVVVLTDKDTSLEYSKNAAT 336
QY 288 TLRNFCANQKLNKVDKHPYWDATLFTRODLCGATCTDTGMADVGMCDPKRSCSV 347
Db 337 TLKNFCKWQHQNQLGDDHEHYDAAILFTREDLCGHSCDITLGMADVGTICSPSCAV 396
QY 348 IEDDGLPSAFATHELGHVFNMPHDNVKVCBEVFKLKRANHMSPILQIDRANPWSACS 407
Db 397 IEDDGLHAFTVAHEIGHLLGLSHDDSKFCEENFTTEDKRLMSILTSIDAKPWSKCT 456
QY 408 AAILTDFLDGSHGCLLDQPSKPTSLPDLPGASYTILSQCELAFGVSKPCPYMOYCTK 467
Db 457 SATITEFLDDGHGNCLLDPRKQTLGPEELPGQTYDATQCNTLTFGPEYSVCPGMDVCAR 516
QY 468 LWCTGKAKGQWCVOTRHPNADGTSCEGKCLKLGACVER-----HNLNHRVDGSKAKW 522
Db 517 LWCVVROGQWVCLTKLPAVEGTPCGKRGVCLQKCKVDTKKYVSTSSH---GNWGSW 573
QY 523 DPGPCSRNTGGGVQOLARROCTNPTPANGGKYCEGVKVRYSNLEPCPSASGSKFREE 582
Db 574 GPMQCSRSRSGGVQOFAYRHCNPNAPRNSGRYCTGRKAIYRSCSVTPCP---PNGKFRHE 631
QY 583 QCEAFNGYNHSTRRLTLAVAWPKYSVSPDRCKCLICRANGTGYFVYLPKPVVDGTLCS 642
Db 632 QCEAKNGYQSDAKGVKTFVEMVPKYAGVLPADVCKLTCRAKGTGYVWFSPKVTDTGECR 691
QY 643 PDSTSVQVQKCKIKAGDGNLKKRFDKCGVCGDNKCKKYTGLTFKPMHGYNFVVAI 702
Db 692 PYSNSVCRGCVRTGCDGLIGSKLQYDKCGVCGDNSSCTKIIGTFNKKSKGYTDVRI 751
QY 703 PAGASSIDIRQYKGLIGDDNVLALNSQKYLNLGHFVVSVAVERDLVVKYGLLRYSGT 762
Db 752 PEGATHIKVRQFRAKQDTPREPAYLALKKTKTEYLINGKYMISFETIIDINGVMYSGW 811
QY 763 GTAVESI-----QASRILEPLVEVLVSKWMPTRVRSFYFLPKPEREDKS----SHPKD 814
Db 812 SHRDDFLUGMYSATKEI-----LIVQLATDPTKALGVRYSFYFPKKTQKRVNSVISHGNS 868
QY 815 PRGFSVLHNSVLNSNOVEQDPRPARVWAGSWGSCGSLGKRAVDCRGSGAGRT 874
Db 869 KVGPI---HSTQL-----QWYTGFWLACSRCTDGTGWHTRVYQCOQ-DGNRKL 909
QY 875 VPACDAAHRRVETQAC 890
Db 910 AKGCLLSQRPSAPKQC 925

RESULT 9
ATSS_HUMAN
ID ATSS_HUMAN STANDARD; PRT; 930 AA.
AC Q9UNAO; Q9UKP2;
DT 16-OCT-2001 (Rel. 40, Created)
```

```
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2)
DE (ADMP-2) (ADAM-TS 11).
GN ADAMTS5 OR ADMP2 OR ADAMTS11.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=99367476; PubMed=10438522;
RA Abbaszade I., Liu R.-O., Yang F., Rosenfeld S.A., Ross O.H.,
RA Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M.,
RA Wynn R., Duke J.L., George H.J., Hillman M.C. Jr., Murphy K.,
RA Wiswall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H.,
RA Ito Y., Newton R.C., Magolda R.L., Trzaskos J.M., Hollis G.F.,
RA Arner E.C., Burn T.C.;
RT "Cloning and characterization of ADAMTS11, an aggrecanase from the
RT ADAMTS family.";
RL J. Biol. Chem. 274:23443-23450(1999).
RN [2]
RS SEQUENCE FROM N.A.
RC MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudo J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsukeyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Deyand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Leirach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [3]
RS SEQUENCE OF 413-930 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=99395124; PubMed=10464288;
RA Hurskainen T.L., Hirohata S., Seidlin M.F., Apte S.S.;
RA "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RA zinc metalloproteases.";
RT J. Biol. Chem. 274:25555-25563(1999).
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN
CC PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
CC site.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVEL, PRIMARILY IN PLACENTA
CC BUT ALSO IN OTHER TISSUES, SUCH AS HEART AND BRAIN, AND ALSO
CC CERVIX, UTERUS, BLADDER, ESOPHAGUS, RIB CARTILAGE,
CC CHONDROBLASTOMA, FIBROUS TISSUE AND JOINT CAPSULE FROM AN
CC ARTHRITIC PATIENT.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF142099; AAD49577.1; -
 DR EMBL: AP001698; BAA95504.1; -
 DR EMBL: AF001697; BAA95503.1; -
 DR EMBL: AF141293; AAF02493.1; -
 DR HSP: Q9PW35; LBUD.
 DR MEROPS; M12.225; -
 DR Genew; HGNC:221; ADAMTS5.
 DR MTM; 605007; -
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR002870; Pep_M12B_propep.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSPL.
 DR InterPro; IPR000130; Zn_MTPetdse.
 DR Pfam; PF00090; tsp.1; 2.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR SMART; SM00209; TSPL; 2.
 DR PROSITE; PS00215; ADAM_MPRO; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS00092; TSPL; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT SIGNAL 1 16
 FT PROPEP 17 261
 FT CHAIN 262 930
 FT SITE 209 209
 FT METAL 410 410
 FT ACT_SITE 411 411
 FT METAL 414 414
 FT METAL 420 420
 FT DOMAIN 485 566
 FT DOMAIN 567 623
 FT DOMAIN 732 874
 FT DOMAIN 875 930
 FT DOMAIN 37 41
 FT DOMAIN 257 261
 FT CARBOHYD 498 498
 FT CARBOHYD 728 728
 FT CARBOHYD 802 802
 FT CARBOHYD 807 807
 FT CONFLICT 138 138
 FT CONFLICT 614 614
 FT CONFLICT 692 692
 SQ SEQUENCE 930 AA; 101715 MW; B64281502F28193B CRC64;
 Query Match 35.8%; Score 1847.5; DB 1; Length 930;
 Best Local Similarity 44.0%; Pred. No. 13e-122;
 Matches 392; Conservative 113; Mismatches 291; Indels 95; Gaps 22;
 QY 48 SGDQGLIFQITAFQEDFYHLTPAQFLAPAFSTHGLVPLQGLTGSSDLR---RCFYS 104
 DB 82 SGGKGVLYVAGGRFLDLDERDGSVGIAGF-----VPAGG--GTSAPWRHRSHCFYR 133
 QY 105 GDVNAEPDPSFAVSLCGGLRCAGFYCAEYVISPL---PNASAPAAQNSQGA---HLL 157
 DB 134 GTVDASPRSLAVFDLCGLDGFFAVKHARYTLKPLLRGPNAEEKRGYDGSARILHYV 193
 QY 158 QRRG-----VPGGPS-----GDPTRCGVAGSWNPAIL---RALDPYKP 193
 DB 194 TREGSFALPRASCETPASTPEAHEHAPAHNSNPGRALAS---QLLDQSALSP--- 246
 QY 194 RRAGFGES---KSRNRGRKRFVSIPIRYVETLVVADESWKFGADLEHYLLLTATAA 250
 DB 247 -AGGSGPQTWVRRRRS-----ISRARQVELLLVADASMARLYGRGLOHYLLTASTAN 299
 QY 251 RLYRHSILNPINIVVKVLLLRDRDSGPKVTGNAALTFRNFCAWQKLNKVS DKHPEY 310

DB 300 RLYSHASTENHIRLAVVYVVLGDKDSLEVSNAATTLKFNCKQWQHOLQDDEEHY 359
 QY 311 DTAILFTRODLGAGTCDTLGMDVGTMCDPKRSCSVIEDDGLP5APFATTAHELGHVFNMP 370
 DB 360 DAAILFTREDLCGHHSDDTLQMDVGVICSPERSCAVIEDDGLHAATVAHEIGHLIGLS 419
 QY 371 HDNVKVCVEEFGKLRANHMSPPTLIQIDRANPWSACAAIITFDLDSGHGDCLLDQPSKP 430
 DB 420 HDDSKFCEETFGSTEDKRLMSILTSIDASKPWSKTSATITEFLDDGHCNCLDLPRQ 479
 QY 431 ISLPEDLPAGSYTLISQCELAFGVSGPCPYMCTYKLTWCTGKAKGOMQVQTRHFWADG 490
 DB 480 ILGPEELPGQTYDATQCNTLTFGPEYSVCEGMVYCARLCAVAVQGMVCLTKKLPAVEG 539
 QY 491 TSCGEGKCLKAGACVER-----HNLNKHVRVDGSAKWDVPGPCSRTCGGVQLARRQCTN 545
 DB 540 TPCGKGRICLOGKCVDTKTKKYYTSSH---GNMGWSGWSGCSRSCGGGVQFAYRHCNN 596
 QY 546 PTPANGGKYCEGVRYKVRNSCNLEPCPSASGKSPREOCEAFNGYNHSTNRLTLAVAWVP 605
 DB 597 PAPRNNGRYCTGKRAIYRSCSLMPCP--PNGKSFHCEQCEAKNGYQSDAKGVKTFVEWVP 654
 QY 606 KYSGVSPDKCKLICRANGTCYFVYVAPKVVDTGLCPDSTSVCGKCIKAGCDGNGLS 665
 DB 655 KYAGVLPADVCKLCTRAKGTGYVYVSPKVDGTECPYSNSVCVRKCVKTCGCGIIGS 714
 QY 666 KRFDKCGVCGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNY 725
 DB 715 KLOYDKCGVCGDNSSCTKIVGTENKSKGYTDVVRIPEGATHIKVQFAKADQTRFAY 774
 QY 726 LALKNSQKYLNGHFVYSAVERDLVVKGLSYSGTGATVESL-----QASRPLEPLT 780
 DB 775 LALKKNGEYLNGKYMISTSETIIDINGTVNYSWMSHRDDFLHGMGYSAKEL---LI 831
 QY 781 VEVLSVGMKTPRVRYSPYLPKEPREDKSSHPKDPGRGSPVLSHNSVLSL-SNQVEQPDPR 839
 DB 832 VOILATDPTKPLDVRYSEFPVK-----STPK-----VNSVTSHGSKNKGSHTSQP 877
 QY 840 PARWVAGSWGPCASGSGLOKRAVDCRGSAQGRTPVACDAHRPVETQAC 890
 DB 878 --QWVTGWLACSRCTDGTGWHTRTVQCQ-DGNRKLAKGCLPSQSPSAFKQC 925
 RESULT 10
 ID ATSA_RAT STANDARD; PRT; 630 AA.
 AC Q9ESP7; Q9ESP8; Q9ESP6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADAMTS-4 precursor (ec 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)
 DE (Fragment).
 DE ADAMTS4.
 GN ADAMTS4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=20415831; PubMed=10961658;
 RA Satoh K., Suzuki N., Yokota H.;
 RT "ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin
 RT motifs) is transcriptionally induced in beta-amyloid treated rat
 RT astrocytes.";
 RL Neurosci. Lett. 289:177-180(2000).
 CC !- FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
 CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
 CC DESTRUCTION OF AGGECAN IN ARTHRITIC DISEASES.
 CC !- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
 CC site.

DR pfam: PF00090; tsp_1; 5.
DR pfam: PF01421; Reprolysin; 1.
DR pfam: PF01562; Pep_X12B_propep; 1.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR Hydrolase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT NON_TER 1
FT PROPEP <1 207 BY SIMILARITY.
FT CHAIN 208 1077 ADAMTS-10.
FT METAL 366 367 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 367 367 BY SIMILARITY.
FT METAL 370 370 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 376 376 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 434 520 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 578 679 DISINTEGRIN-LIKE.
FT DOMAIN 680 802 CYS-RICH.
FT DOMAIN 521 577 SPACER.
FT DOMAIN 799 860 TSP TYPE-1 1.
FT DOMAIN 862 918 TSP TYPE-1 2.
FT DOMAIN 922 976 TSP TYPE-1 3.
FT DOMAIN 981 1031 TSP TYPE-1 4.
FT CARBOHYD 64 64 TSP TYPE-1 5.
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 769 769 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1077 AA; 118072 MW; 3914DE18DCBFF587 CRC64;

Query Match 25.1%; Score 1296; DB 1; Length 1077;
Best Local Similarity 35.4%; Pred No. 1.3e-83;
Matches 335; Conservative 112; Mismatches 376; Indels 123; Gaps 32;

QY 24 EVVVPRLDPDINGR-----RYWRGPDGSGDGLIFQITAFQEDFYHLHPDAQF 74
DB 14 EIAFPTRVDRH-NGALLAFPPPPRRQRGTGATAESRLFYKVASPSHFLLMLTRSSRL 71
QY 75 LAPAFSTEHLGVPLQGLTGSSDLRCFYSGDYNAE-PSQFAVSLCGGLRGAFGYRGAE 133
DB 72 LAGHVSVEYW--TREGIAWQARAPRCHLYAGHQQGASSHVAISCTCGGLHGLIVADEEE 129
QY 134 YVISPPLNASAPAAQRNSQGAHLQRQVPGSPGSDPTSRGV--ASGW--NPAILRALD 189
DB 130 YLIEPLHGGPKGSRSPESGPHVYKRSLRHPHLD--TAGGYRDEKPKWGRPWMLRTLK 187
QY 190 PYKPRRAGFGESSRRRSRAKRFVPIPRVETLVVADESQWVFKHG-ADLEHYLLLLAT 248
DB 188 P--PPARPLGNETERGQPG-LKRSVSRERYETLVVADKMMVAYHGRDVEQYVLAIMI 244
QY 249 AARLYRHSILNPINIVVKKVLLLRDSDGPKVTGNAALFLRNFCAWOKKL-----NK 301
DB 245 VAKLFQSSLSGTVNIIIVTLRLITLEDQPLETHIAGKSLDSFCWKQKSVNHSCHGNA 304
QY 302 VSDRHPYETDAILFTQDILC--GATTCDTLGMADVGTMDPKRSCSVIEDDGLPSAFT 359
DB 305 IPENGVAHDTAVLITDYICVKNKPCGTGLGLAPVGGMCEERSCSVNEDIGLPOAFTI 364
QY 360 AHELGHVFNPHDNVKKVEEVFG-----KLRANHMSFTLIQIDRANP--WSACSAALITD 413
DB 365 AHEIGHTFGMNHDCVGNSSCGARGQDPAKLMAAHTMTK-----NPFVWSSCNRDYITS 417
QY 414 FLDGSGDCLLDQPSK-PISLPEDLPKASVYTLSSQCEFLAFGVSGKPCPYMYQYCTKLWCTG 472
DB 418 FLDGSLGLCLNRRPPRDFVYTPAPQADYADACQRFQHGKVSQCKYGEVCSLWCLS 477
QY 473 KAKQMVQCQTRHFPWADTSGEGKL-----CLKGACVVERHNLKRVGDSNAKWDPYGPC 528
DB 478 KSNR---CITNSIPAAGETLCQHTIDKGMCKYKRCVCP-FGSRPEGVGDGAWPWTWGD 533

QY 529 SRTCGGVQLARROCTNPTPANGKGYCEGVVYKRSCLNLEPCPSSASGSKSFREUQCEAFN 588
DB 534 SRTCGGVSSSRHCDSPRTIGGKYCLGERRRHRSCNTDDCPGS--QDFREVQCSEFD 591
QY 589 G-----YNHSTNRLTLAVAWVPKYSYGVSPDKCKLICRANGTGYFVVLAPKVVDTGLC 641
DB 592 SIPPRGFYKWKTYR-----GGGVK---ACSLTSLAEGFNFTYTERAAAVVDGTPC 638
QY 642 SPDSTSVQVQCKITKAGCDGNLGSKKRFRDGCVGCGDNKSKKVTGLFT--KPMHYNYFV 699
DB 639 RPDVDICVSGECKHVGCDRVLGSLREDKRCVCGDGSACETIEGVSPSPAGYEDV 698
QY 700 VAIPAGASSIDIIRQGYKGLIGDDNYLALKNQSKYLLNHFVFAVERDLVVK:SLRLY 759
DB 699 VWIPKGVSHFIQD-----LNLSLHLALKDQESLLEGLPTGPQPHR-LPLAGTTFQL 752
QY 760 SGTGTAVESLQASRPILEPLTVLSVGMTPPRVRSFYLPKPREDKSSHEKDPGRPS 819
DB 753 RQGPQVQSLEALGPTNASLIVVLARTEL--PALRYRFNAP----- 792
QY 820 VLHNSVLSLNQVBEQDPRPPARVAGSWGPCSGSLQKRAVDCRGSGAGRTV-PAC 878
DB 793 -----IARDSLPPYSWHYAPWTKCSAQCAAGSGVQVAECRNQDSSAVAPHY 839
QY 879 DAHR--PVEQACG-EPC-PTWELSAWSPCSKSGRGRFORSLK 920
DB 840 CSAHSLKLPKQACNETPCPDWVWVGNWLSRSCDAGVRSRVSVC 885
RESULT 12
AT12_HUMAN
ID AT12_HUMAN STANDARD; PRT; 1593 AA.
AC PS8397;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-DE TS12).
DE GN ADAMTS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=21264577; PubMed=11279086;
RA Cal S., Arqueelles J.M., Fernandez P.L., Lopez-Otin C.;
RT "Identification, characterization, and intracellular processing of ADAM-TS12, a novel human disintegrin with a complex structural organization involving multiple thrombospondin-1 repeats.";
RL J. Biol. Chem. 276:17932-17940(2001).
CC CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC CC -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely expressed in gastric carcinomas and in cancer cells of diverse origin.
CC CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY SIMILARITY).
CC CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC CC -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1 DOMAINS.
CC CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC CC -!- SIMILARITY: CONTAINS 8 TSP TYPE-1 DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

(Procollagen I/II amino-propeptide processing enzyme).
ADAMTS2 OR PCNP OR PCNP1.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS LPNP1 AND SPNP1), AND DISEASE.
TISSUE-Skin; PubMed=10417273;
MEDLINE=99347935; Li S.-W., Schwarze U., Petty E.,
Collige A., Sieron A.L., Krakow D., Cohn D.H., Reardon W.,
Wertelecki W., Wilcox W., Prockop D.J., Nusgens B.V.,
Beyers P.H., Lapiere C.M., Prockop D.J., Nusgens B.V.,
"Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis
are caused by mutations in the procollagen I N-proteinase gene";
Am. J. Hum. Genet. 65:308-317(1999).
CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
to fibril assembly. Does not act on type III collagen. May also
play a role in development that is independent of its role in
collagen biosynthesis.
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
alpha-1(I) at Pro-1-Gln and of alpha-1(II) and alpha-2(I) chains
at Ala-1-Gln.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBUNIT: MAY BELONG TO A MULTIMERIC COMPLEX. BINDS SPECIFICALLY TO
COLLAGEN TYPE XIV (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: LPNP1 (SHOWN HERE) AND SPNP1;
ARE PRODUCED BY ALTERNATIVE SPLICING. SPNP1 HAS NO SIGNIFICANT N-
PROCOLLAGEN PEPTIDASE ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVEL IN SKIN, BONE, TENDON
AND AORTA AND AT LOW LEVELS IN THYMUS AND BRAIN.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
SIMILARITY).
CC -!- DISEASE: Defects in ADAMTS2 are the cause of Ehlers-Danlos
syndrome type VIIC (EDS-VIIC), a recessively inherited connective-
tissue disorder characterized clinically by severe skin fragility
and joint hypermobility and biochemically by the presence in skin
of procollagen incompletely processed at the N-terminus.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: AJ003125; CAA03680.1;
MEROPS: M12.301;
Genew: HGNC:218; ADAMTS2.
MTM: 604539;
MTM: 225410;
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_MTPeptase.
DR Pfam: PF00050; tsp_1; 4.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR SMART: SM00209; TSP1; 4.
DR PROSITE: PS00215; ADAM_MEPPO; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE NEG.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.

Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Collagen degradation; Extracellular matrix; Heparin-binding;
Alternative splicing; Ehlers-Danlos syndrome.
POTENTIAL.
FT SIGNAL 1 29
FT PROPEP 30 253
FT CHAIN 254 1211
FT METAL 408 408
FT ACT_SITE 409 409
FT METAL 412 412
FT METAL 418 418
FT DOMAIN 480 560
FT DOMAIN 561 617
FT DOMAIN 618 722
FT SITE 691 693
FT DOMAIN 723 851
FT DOMAIN 852 911
FT DOMAIN 912 974
FT DOMAIN 975 1030
FT DOMAIN 140 43
FT DOMAIN 185 188
FT CARBOHYD 112 112
FT CARBOHYD 251 251
FT CARBOHYD 949 949
FT CARBOHYD 993 993
FT CARBOHYD 1031 1031
FT CARBOHYD 1098 1098
FT CARBOHYD 1145 1145
FT CARBOHYD 1150 1150
FT VARSPLIC 544 566
FT VARSPLIC 567 1211
FT SEQUENCE 1211 AA; 134722 MW; BCEBEF25C23CAD2D CRC64;
SQ
Query Match 23.8%; Score 1231; DB 1; Length 1211;
Best Local Similarity 32.9%; Pred. No. 5.5e-79;
Matches 345; Conservative 136; Mismatches 399; Indels 168; Gaps 47;
QY 14 RTAGSEPP-----EREVVPPIRLDPD-----INGR-----YY 41
DB 38 RLAAADPPGGPLGHGAERILAVEPRTDAGRLVSHVSVSAATSRAGVARRAAFRTPSF 97
QY 42 WRGSDSGDQGLIFQITAFQDFYHLTPDAQFLAPAFSTEHLGVLQGLTGGSSDLRRC 101
DB 98 PGNEEPGSHLFYVIVFGRLHLRLRNARLAPCATMEWQGE--KGTRVEPLLGSC 155
QY 102 FYSGDVN--AEPDSFAAVGCGGLRGAFYGAEBYVTSPLPNASAPAAQNSCG-AHLLQ 158
DB 156 LYVDVAGLAESS-VALSNCGLAGLIRMEEEFFTEPLEKGL--AAQAEKGRHVY 212
QY 159 RRGVPGSGDPTSRGCVASGNPAILRALDPYKPRAGFGESRRSRRAKRFVSIPR 218
DB 213 RRPPTSPLGPPQA---LDTG---ASLSDLSLS-RALGVLEEHANSRRARRHAADD 265
QY 219 Y-VETLVVADESVMKFGAD-LEHYLLTLATAARLYRHPISILNPINIVVVKVLLLRDR 276
DB 266 YNIEVLLGVDDSVQFPGKEHVQKYLTLNINVEIYHDSGAHINVLVRIILLSYGK 325
QY 277 SGPKV-TGNAALTLRNFCWOKKLINVKSDKHPFWDFTAILFTRODLGATCTDLGNADY 335
DB 326 SMSLIEIGNPSQSLENCVRWAYLQOKPDTGHDEYHDHAIFLTQDF-GPSGMQ--GYAP 382
QY 336 GTMCDPKSCSVIEDDGLPSAFTTAHELGHVFNPHDNV-KVCEVEFGKLRANHMSPTL 394
DB 383 TGMCHPVRSCTLNHEDGFSFAFVAHETGHVLMGHEHGDQGNRCGD---EVRLSIMAPLV 439
QY 395 IQIDRANPWSACSAIIITFDLSDHGCDLLDQP---SKPISLPEDLPASVTLISQOCELA 451
DB 440 QAAFRHFHNSRCSQELSRYLHS--YDCLLDDPFAHDWP-ALPQ-LPGLHYSMNEQCRFD 495
QY 452 FVGSGKPCPYMOY---CTKLMCTGAKQWYQVQTRHFPWADGTSCGEGKCLKACV--- 505
DB 496 FGLGYMMCTAFRTFDPCQKLMCS-HPDNPYFCFKKGPPLDGTMCAPGKHCFCGICWLT 554

QY 506 ----ERHNLKRVDPGSGWAKWDYPCGSRCTCGGQVQLARRQCTNPTPANGKKYCEGVRYK 561
 Db 555 PDILKR-----DGSWAWSPGSCSRCTCGTGVKFTQCNDPMPHAGRTCSGLAYD 606
 QY 562 YRSCNLEPCSSAGSKSFREOCFAFNQY-----NHSTNRJTLAVAWPKYSGVSPRD 614
 Db 607 FOLCSRODCPDLSLA--DFREOCQWDLYPEHGDAQHH-----WLP-HEHRDAKE 653
 QY 615 KCKLICRANGTGYPIVLAPKVDGTLGS-PDSTSVCOGKCIKAGCDGNLGSKKRFDKCG 673
 Db 654 RCHLYCESRETGEVWSMKRMVHDGTRCSYKDAFSLCVRGDCRVGCDGVIGSSKQEDKCG 713
 QY 674 VCGDNKSKCVKTLGFTK--PMHGYNEFVAIPAGASSIDIRQYKGLIGDDNYLALKNS 731
 Db 714 VCGDNHCKVKGTFTRSPKKHGYKMFEPAGARHLLQEQ-----VDATSHLAVKNL 768
 QY 732 Q-GKYLNGHFVYSAVERDLVVKGSLRLRYSGTGTAVESLOASRPILFPLTFEVLVSVKMT 790
 Db 769 ETGKFIINEENDVASSKTFIAMGVWEYRDE-DGRETLQTMGLHCTITIVLIPVGD-T 826
 QY 791 PPRVRSFYLPKEP-----RED-----KSSHPKDPRG----- 817
 Db 827 RVSLTYKMHEDSLNVDNNVLEEDSVYEWALKKWSKPCSGCGSQFTKYGCRRRLD 886
 QY 818 PSVLHN---SVLSLSNQVEQ---PDRPPARWVAGSWGPCSASCG-SGLQKRAVDCRGA 870
 Db 887 HKMVHRGFCAALSKPAIRACNPQCSQPVVVTGEWPCSTCGRTGMQVRSVRCIQPL 946
 QY 871 GQRTVPACDAAH---RPVETOACG-EPGP-TWELSAWPCSKSCGRGQRSLKCVG 924
 Db 947 HDNTRSVHAKHNDARPSRRACSRRELCPGRWRAGFWSCSVTCGNGTOERPVP 1002
 QY 925 GRLLANDQCNL---HRKPQELDFCVLRPC 950
 Db 1003 --RTADDSFGICQERPERPARTCRGLGPC 1028

RESULT 14
 AT57_HUMAN
 ID AT57_HUMAN STANDARD; PRT; 997 AA.
 AC Q9UPK4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADAMTS-7 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).
 GN ADAMTS7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP
 RX MEDLINE-99395124; PubMed-10464288;
 RA Hurskainen T.L., Hironaka S., Seldin M.F., Apte S.S.;
 RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, Novel Members of a New Family of Zinc Metalloproteases";
 RL J. Biol. Chem. 274:25555-25563(1999).
 CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PESTIDASE FAMILY M12B.
 CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF140675; AAD56358.1; -
 DR HSSP; P15167; IATL.
 DR MEROPS; M12.231; -
 DR Genew; HGNC:223; ADAMTS7.
 DR MIM; 605009; -
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR002870; Pep_M12B-propep.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR000130; Zn_Mtpeptidse.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF01562; Pep_M12B-propep; 1.
 DR SMART; SM00209; TSP1; 2.
 DR PROSITE; PS00142; ZINC-PROTEASE; 1.
 DR PROSITE; PS0215; ADAM_MEPRO; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT SIGNAL; 1 27 POTENTIAL.
 FT PROPEP; 28 232 BY SIMILARITY.
 FT CHAIN; 233 997 ADAMTS-7.
 FT SITE; 204 204 CYSTEINE SWITCH (POTENTIAL).
 FT METAL; 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE; 389 389 BY SIMILARITY.
 FT METAL; 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL; 398 398 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DOMAIN; 462 537 DISINTEGRIN-LIKE.
 FT DOMAIN; 538 594 TSP TYPE-1 1.
 FT DOMAIN; 595 697 CYS-RICH.
 FT DOMAIN; 698 914 SPACER.
 FT DOMAIN; 915 990 TSP TYPE-1 2.
 FT CARBOHYD; 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD; 693 693 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD; 778 778 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 997 AA; 6587044ED02FC104 CRC64;
 Query Match 23.5%; Score 1215.5; DB 1; Length 997;
 Best Local Similarity 35.1%; Pred. No. 5.4e-78;
 Matches 353; Conservative 117; Mismatches 376; Indels 159; Gaps 46;
 QY 1 MLLGLILTAFLAPGRTAGGSEPER---EVVVPRLDP-----DINGRR- 39
 Db 18 LLLCLALAPGAPGAPGRATEGRAALDIVHPVRVDAGGSLSYELWPRALKKPDVSVRD 77
 QY 40 ---YY---WRGPEDSGDGLIFQITAFQEDFYHLHTPDQFLAPAF--STEHLGVPLQGL 91
 Db 78 APAYELOYRGRE-----LRFNLTAQ-----LAPGFVSETRREG---GL 116
 QY 92 TGGSSDLR---RCFYSGDVNABPD---SFAAVSLCGSLRGAGCYRGAEEVISPPLNASA 144
 Db 117 --GRAHTRAHTPACHLLGEVQ--DELEGGLAAISACDGLKGVFQLSNEDYFIQPLD--SA 171
 QY 145 PAAGRNQGAHLQRRGVP-----GGPSGPTSCGVASGNPAILRALDPYKPRRAGFG 199
 Db 172 PARPGHAQ--PHVYKQAPERLAQRGDSSAP-STCGV-----QVYPELESRRERW 219
 QY 200 ESRSRRSRGRKRF----VSIPRYVELTVVADESWKVFKHG--ADLEHYLLTLLTATAARLYR 254
 Db 220 EQROOWRRPLRLRHORSVSKWKVETLVVADAKMVEYHGQPVQSVYLTIMNVAQLPH 279
 QY 255 HPSILNPINIVYKVLRLDRDSGPKVTGNAALTLRNFCAMQKLNKVSQKHFYEDWDTAI 314
 Db 280 DPSIGNPIHITIVRLVLEDEEDLKITHDADNTLKSFKWKQKSNMKGDAPLHDDTAI 339
 QY 315 LFTQDLCGATT--CDTLGMADVGTMCDPKRCGSVIEDDGLPSAFTTAHELGHVFNPHD 372

Db 340 LLTKDLCAAMNRPCEITGLSHVAGMCQPHRSCSINEDTGLPLAFTVAHELGHSGFQIHD 399
Qy 373 -NVKCEVFQKLRANHMSTLIDRANPWSACSAIITDFLDSCHGCLLDQSKP- 430
Db 400 GSGNDCEPV-GK--RPFIMSPQLLYDAAPLTWRSRCSQYITRFLDKWGLCLDDPPAKDI 456
Qy 431 ISLPEDLPASYSYLSQCELAFGVSGPCPYM-QYCTKLWCTGKAKGMVQCOTRHPFPWAD 489
Db 457 IDFPSPVPPGVLYDVSHQCRLOYGAYSACEDMDNVCHTLWCSVGT----TCHSKLDAVD 512
Qy 490 GTSCEGKGLCKLGACVERHNLKHKRVDSWAKWDPYGPCSRTOCGGVOLARRQCTNTPPA 549
Db 513 GTRCGENKWLSCGECVP-VGFRPEAVDGGWSGSAWSSICRSRCSGMSVQSAERQCTQTPK 571
Qy 550 NGGKYCEGVRYKRSCLNEPCPSSASGSKSPREEOCEAFNGYNHSTNRLTLAVANVPKYS 609
Db 572 YGGRYCVGERKFRCLNLOQCP--AGRPSFRHVQCSHFDAWYKGLQHT----WPPVVD 625
Qy 610 VSPRDKKLCIRANGTYFYVLAPKVVDTGLCSP--DSTSVGVQKCIKAGCDGNLGSKK 667
Db 626 VNP---CELHCPANEYAKKLRAVDVDTGTCYQVRSRDLCLINGICKNVGCDFEIDSGA 682
Qy 668 RFDKCVCGGDNKKVYGLTFRPMH-GYNFVVAIPAGASSIDIFORGYKGLIGDDNLYL 726
Db 683 MEDRGVCHGNGSTCHTYSGTFFEEAGLGYVDVGLIPAGAREIRIQE-----VAEAAANFL 737
Qy 727 ALKNSQ-GKYLNGHFVYSAVERDLVVKGSLLRYSGTGTAVESLQASRPILPLETVEVL- 784
Db 738 ALRSEDPEKIFLNGWTIQ--WNGDQVQAGTFTYARRGN-WENLSPGTPKEPWIQVPA 795
Qy 785 -----SVGKMTTPRVYSFYLPKPREDKSSHFKDPRGFSVLHNSVLNSNQVEQDD 837
Db 796 SRCPGGSGRGVPRSTLHGRSRPGVSPGVTGEPGPPAAASTSVS----- 845
Qy 838 RPARW---VA-----GSGWPCASCGSLQKRAVDCRS-----AGQRTVPACDAHRP 884
Db 846 -PSLKWPNLVAHVHGGWG--QAPLGLGWRHRLVLMGPRFLTQLLFOESNPGVHYEY-T 901
Qy 885 VETOACGE-----PCTWELSAWSPCKSCSGRG--FORRSILKCVG 922
Db 902 IHREAGGHDEVPFPVSHFHYGWTCTVTCGKEGWRHSPTCRG 946

RESULT 15
ATS2_BOVIN STANDARD; PRT; 1205 AA.
AC P79331;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-2 precursor (BC 3.4.24.14) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2)
DE (Procollagen I/II amino-propeptide processing enzyme) (Procollagen I
DE N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (pNPI).
GN ADAMTS2 OR NPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=97225960; PubMed=9122202;
RA Collige A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J.,
RA Lapiere C.M.;
RT "CDNA cloning and expression of bovine procollagen I N-proteinase: a
RT new member of the superfamily of zinc-metalloproteinases with binding
RT sites for cells and other matrix components.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
RN [2]
RP PARTIAL SEQUENCE.

RX MEDLINE=95348096; PubMed=7622483;
RA Collige A., Beschlin A., Samyn B., Goebels Y., Van Beeumen J.,
RA Nusgens B.V., Lapiere C.M.;
RT "Characterization and partial amino acid sequencing of a 107-kDa
RT procollagen I N-proteinase purified by affinity chromatography on
RT immobilized type XIV collagen.";
RL J. Biol. Chem. 270:16724-16730(1995).
CC -!- FUNCTION: CLEAVES THE PROPEPTIDES OF TYPE I AND II COLLAGEN PRIOR
CC TO FIBRIL ASSEMBLY. DOES NOT ACT ON TYPE III COLLAGEN. MAY ALSO
CC PLAY A ROLE IN DEVELOPMENT THAT IS INDEPENDANT OF ITS ROLE IN
CC COLLAGEN BIOSYNTHESIS.
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC alpha-1(I) at Pro-1-Gln and of alpha-1(II) and alpha-2(I) chains
CC at Ala-1-Gln.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBUNIT: MAY BELONG TO A MULTIMERIC COMPLEX. BINDS SPECIFICALLY TO
CC COLLAGEN TYPE XIV.
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: ENZYMIC ACTIVITY IS DETECTED AT HIGH LEVEL
CC IN ALL TYPE I COLLAGEN-RICH TISSUES SUCH AS SKIN, BONES, TENDONS
CC AND AORTA AND AT LOW LEVEL IN BRAIN AND THYMUS. THE MRNA LEVELS
CC WERE DISPROPORTIONATELY HIGH IN HEART, LIVER, RETINA AND MUSCLE.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- DISEASE: DEFECTS IN ADAMTS2 ARE THE CAUSE OF DERMATOSPARAXIS, A
CC RECESSIVELY INHERITED DISORDER CHARACTERIZED BY SEVERE SKIN
CC FRAGILITY AND BIOCHEMICALLY BY THE PRESENCE IN SKIN OF PROCOLLAGEN
CC INCOMPLETELY PROCESSED AT THE AMINO TERMINUS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X96389; CAA65253.1; -
CC MEROPS; M12.301; -
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Repolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_Mrpeptdse.
DR Pfam: PF000090; tsp_1; 4.
DR Pfam: PF01421; Repolysin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR SMART: SM00209; TSP1; 4.
DR PROSITE: PS50215; ADAM_MEROPS; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Collagen degradation; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 253 BY SIMILARITY.
FT CHAIN 254 1205 ADAMTS-2
FT METAL 402 402 ZINC (CATALYTIC) (POTENTIAL).
FT ACT_SITE 403 403 BY SIMILARITY.
FT METAL 406 406 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 474 554 DISINTEGRIN-LIKE.
FT DOMAIN 555 611 TSP TYPE-1 1.
FT DOMAIN 612 716 CYS-RICH.
FT DOMAIN 717 845 SPACER.

```
FT DOMAIN 846 905 TSP TYPE-1 2.
FT DOMAIN 906 968 TSP TYPE-1 3.
FT DOMAIN 969 1024 TSP TYPE-1 4.
FT SITE 685 687 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 31 35 POLY-ALA.
FT DOMAIN 177 180 POLY-GLU.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 943 943 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 987 987 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1092 1092 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1144 1144 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1205 AA; 133887 MW; 7B5B2324A5320371 CRC64;

Query Match 23.1%; Score 1192; DB 1; Length 1205;
Best Local Similarity 31.7%; Pred. No. 3.1e-76;
Matches 329; Conservative 134; Mismatches 403; Indels 172; Gaps 42;

Qy 1 MLLGLTLTAPAGTAG-----GSEPEREVVVPIRLD----- 32
Db LLLLLLLPADARLAAAAADPPGPGQGHGAERILAVPVRTDAQRLVSHVSNATAPAG 75
Qy 33 -----PDINGRRYWRGPDSDGDOGLIFQITAFQEDFYHLHTPDQFLAPAFST 81
Db 76 VRTTRAAPAQIPGLSG-----GSEEDPGGRLEFYNVTFGRDLHLRLRPNARLVAPGATV 129
Qy 82 EHLGVPLQGLTGGSSDLRRRCFYSGDV-NABPDSPAFAVSLCGGLRGAGFYGAEEVVISPLP 140
Db 130 EWQGE--SGATRVFPLGLTCLYGVGVAGLAESSVALSCDGLAGLIRMEEEFFIEPLE 187
Qy 141 NASAPAAQRNSQG-AHLLORRGVP-----GGPSGDPTSRGVSAGSNPAIRALDPYKP 193
Db 188 KGL--AAKEAEQGRVHVYHRTPTSRPPLGGPOALDT---GISADSLSLRAL----- 237
Qy 194 RRAFGESRRSRRAKRFVSTPRY-VELLVVADESVMVFXHAD-LEHYLLTLLATAAR 251
Db 238 ---GVLEERVNSSRRRRRAADDDYNIIEVLGVDDSVQFHGTEHVQKYLTLMLNIVNE 294
Qy 252 LYRHPSTILNPINIVVVLRLORDSGPKV-TGNAALTIRNFCANOKKLNKVSDDKHPYEW 310
Db 295 IYHDESIGAHINIVVVLRIILLSYKGSMSLIEIGNPSOLENVCRWAYLOQKPDTHDEYH 354
Qy 311 DTALFTRODLCGATTCDTLGMADYGTMCDPKRSVIEDDGLPSAFTTAHELGHVFNMP 370
Db 355 DHAIFLTRQDF-GPSGMQ--GYAPVTGCHPVRSCTILNHEDGFSSAFVAHETGHVLGME 411
Qy 371 HDNV-KVCEEVFGKLRANHMSPTLIOIDRANPWSACSAIITDFLDSHGDCLLDQP-- 427
Db 412 HDGQGNRCGD---EVRIGSIAPLIVQAARFHWRSRCSQOELSRYLHS--YDCLRDDPFT 466
Qy 428 -SKPISLPEDLPGASVYTLSSQOCELAFGVGSKPCPYMQY---CTKLACTGKAKQMVQCTR 483
Db 467 HDWP-ALPQ-LPGLHYSNNEQCRDFGLGYMMCTAFRTDPCKQLWCS-HPDNPYFCKTK 523
Qy 484 HFPWADTSCGEGKCLKLGACV-----ERHNLNKHVRVDGSAKWDPYGPCSRTCGGV 536
Db 524 KGPPLDCTMCAPGKHCFCGHCILWTPDLKR-----DGNWGANSPFGSCSRTCTGTG 575
Qy 537 QLARRQCTNTPANGKYCGVRVYKVRSCNLEPCPSASGKSFPREECCAFNGY----- 590
Db 576 KFRTRQCDNPHANGGRTCSGLAYDQFCNSQDCPDALA--DFREECCQWDLYFEHGDA 633
Qy 591 -NHSTNRLTLAVAVPKYSGVSPRDKCKLICRANGTCGYFYVLAPKVVDGTLCG-PDSTSV 648
Db 634 QHH-----WLP-HEHRDAKERCHLYCESKETCEVYVSMKRMVHDGTRCSYKDAFSL 682
Qy 649 CVQCKICKAGCDGNLGSKRKRFKDCGCGGDNKSKKVTGLTKPMH--GYNFWVAIPAGA 706
Db 683 CVRGDCRKCVDGVISSKQEDKCGCGGDNHCKVVKVGTFSRSPKGLGYIKMFEIPAGA 742
```

```
QY 707 SSTDIRQYKGLIGDNDYLALKNQO-GKYLINGHFVWSAVERDLVVKGLLRYSGTGTA 765
Db 743 RHLLIQEAD-----TTSHLAVKNLETKFILNEENDVPNSKTFITAMGVWEYRDE-DG 796
QY 766 VESLOASRPILLEPLTVEVLVSVGKMTPPRYRYSFYLPKEPREDKSSHPKDPRGFSVLHNSV 825
Db 797 RETLQTMGPLHGTTITVLVIPEG---DARISLTY-----KYMIHEDS 834
QY 826 LSLSNQVEQPDPRPPARWVAGSWGSPCSAGSGLOKRAVDRCGSAGQRTVPA--CDAAH 883
Db 835 LNVDDNNVLEDDSVGVENALKKWSPCKSCGGSGQFTKYGCRRRLDHKMWHRGFCOSVSK 894
QY 884 P-----VETOAGCEPCTWELSAWSPCKSCGR-GFQRRSLKCV-----GHGGRLLARDQ 932
Db 895 PRAIRRTCNPQECQ--PVMVTGEWEPESRSCGRTGMQVRSVRCVOPLHNNITRSVHTKH 952
QY 933 CNLHRKPQELDFCVLRPC 950
Db 953 CN-DARPEGRACNREL 969
```

Search completed: May 9, 2003, 15:22:42
Job time : 22 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:20:56 ; Search time 27 Seconds
(without alignments)
3382.507 Million cell updates/sec

Title: US-09-965-631-4
Perfect score: 5162
Sequence: 1 MLLGLITLAFAGRTAGGSE.....DQNLHRRKPOELDFCVLRPC 950
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2485.5	48.1	951	2 T00017	gene ADAMTS-1 prot
2	1913	37.1	837	2 T00355	hypothetical prote
3	1604	31.1	550	2 T47158	hypothetical prote
4	1587	30.7	2165	2 T21371	hypothetical prote
5	1192	23.1	1205	2 T18517	procollagen N-endo
6	681	13.2	1558	2 C89114	protein C37C3.6a
7	681	13.2	2167	2 T34395	hypothetical prote
8	559.5	10.8	1444	2 T18856	angiogenesis inhib
9	510.5	9.9	860	2 T16892	hypothetical prote
10	493	9.6	951	2 T00260	hypothetical prote
11	450.5	8.7	957	2 T15976	hypothetical prote
12	411.5	8.0	1059	2 T22545	hypothetical prote
13	343.5	6.7	571	2 S24789	jararagin C precu
14	340.5	6.6	609	2 S55270	catrocollastatin p
15	324	6.3	903	2 S60257	meltrin alpha - mo
16	323.5	6.3	617	2 S48160	metalloproteinase
17	320	6.2	826	2 A60385	monocyte surface a
18	313	6.1	549	2 S48169	metalloproteinase
19	311	6.0	411	1 HYSNFA	fibrolase (EC 3.4.
20	308	6.0	407	2 S66260	metalloproteinase
21	307	5.9	789	2 S28259	androgen-regulated
22	304.5	5.9	616	2 A55796	ecarin precursor -
23	298.5	5.8	480	1 A30065	trigramin precurs
24	292	5.7	478	2 JC4880	fibriolytic metal
25	291	5.6	414	2 J41609	atrolysin C (EC 3.
26	290.5	5.6	481	2 JC4342	fibriolytic prote
27	289.5	5.6	610	2 JC7530	vascular apoptosis
28	286	5.5	414	1 HYSAC	atrolysin C (EC 3.
29	284	5.5	481	2 S43125	trimucin precursor

30	277.5	5.4	814	2 G02390	disintegrin-like m
31	277	5.4	478	2 A43296	atrolysin E (EC 3.
32	267	5.2	414	2 S41608	atrolysin B (EC 3.
33	265	5.1	670	2 I65967	disintegrin-like m
34	256.5	5.0	1074	2 JC5928	semaphorin F precu
35	256	5.0	776	2 S28258	androgen-regulated
36	255.5	4.9	1584	2 T00026	brain-specific ang
37	253.5	4.9	952	2 T18900	disintegrin and me
38	251	4.9	1042	2 T26644	hypothetical prote
39	248.5	4.8	1170	1 TSHUP1	thrombospondin 1 p
40	247	4.8	478	2 JQ1301	hemorrhagic protel
41	245.5	4.8	1170	2 A40558	thrombospondin 1 p
42	244.5	4.7	419	2 A59414	metalloproteinase
43	236.5	4.6	823	2 S18968	cytostatin precurs
44	234.5	4.5	508	2 T22836	hypothetical prote
45	234	4.5	419	2 S41607	atrolysin A (EC 3.

ALIGNMENTS

RESULT 1

T00017

gene ADAMTS-1 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000

C:Accession: T00017

R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.

Genomics 46, 466-471, 1997

A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1

A:Reference number: Z14055; MUID:98110583; PMID:9441751

A:Accession: T00017

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-951 <KUN>

A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057

A:Experimental source: strain 129SVJ

C:Genetics:

A:Gene: ADAMTS-1

A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2

C:Superfamily: thrombospondin type 1 repeat homology

F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 48.1%; Score 2485.5; DB 2; Length 951;
Best Local Similarity 48.0%; Pred. No. 2,7e-161;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

QY 1 MLLGLITLAFAGRTAGG--SEPERVVVPIRLDPDINGRRYYWRGP-EDSGDGLFQI 57

Db 20 LLLASITMLLCARGAHRPTDEELVLP-SLE-----RAPHGDSITTRL--RL 66

QY 58 TATQDEFYHLTPDAQFLAPAFSTEHLGVPLQGLTGS-----SDLRCYSGDVNA 109

Db 67 DAFQQLHLQPDGSLFGFTQTQV-----GRSPGEAQHLDPDGLAHCFYSTGNG 121

QY 110 EPDSFAAVSLCGLRGAFGRGAEYVISPUNAS-----APAAQRNSQGA----HLQRRG 161

Db 122 DPGSAALSCEGVRGAFYLOGEEFFIQAPGVATERLAPVAPPEESSARPQHILRRR- 180

QY 162 VPGGPGSDPTSRGCVASGMNPAILRALDPYKPRRAGFGESESRRR----- 206

Db 181 ----RRSGGAKCGVMD-----DETLPISDSRPESQNTNRNPVDRPTPODAGRP 226

QY 207 ----SGRAKRFVSIPIRYVETLVVADESVMKFGCADLEHYLLTLLATAARLYRHPSLNFI 262

Db 227 SGPGSIRKKRFVSIPRYVETMLVADQSMADFHGSLGLKHYLLTFLSVAARFYKHPSTRNST 286

QY 263 NIWVKVLLLRDSDGPKVTGNAAALTNRNFCWAKKLNKVSOKHPPEYWTAILFTRQDLG 322

Db 287 SLVVVKILVIYEQKGPVEVTSNAAALTNRNFCWAKKLNKVSOKHPPEYWTAILFTRQDLG 346

QY 323 GATTCDTLGMADVGTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEFG 382

Db 323 GATTCDTLGMADVGTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEFG 382

```
Db 347 GSHCTDLGMADVGVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCAASLNG 406
QY 383 KLRANHMSPTLIQIDRANPWSACSAAIITDLDGSHGDCLLDQSPKIPSLPEDLPGASY 442
Db 407 VTGDSHLMASMLSSLDHSQWPSCSAYMYTSFLDNGHGECLMDKPNPIKLPSDPLGTYL 466
QY 443 TJSQOCELAFGVGSKPCP-YMYQCYTKLWCTGAKAKOMVQOTRHFPWADGTSCEGKGLCKL 501
Db 467 DANROQFTFGESKCHPCDAASTCTTLWCTGTSGGLLVQCTKHFPWADGTSCEGKGCVS 526
QY 502 GACVERHNLNK---RVDSWAKWDPYGPCSRTGCGVQLARROCTNPTBANGKCYCEGV 558
Db 527 GKCVNKTDM-KHFATPVHGSWGPWPGDCSRTGCGGVQVYTMRECDNPVENKNGKCYCEGK 585
QY 559 RVKYSRNLPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVARVPKYSVGVSPDRCKKL 618
Db 586 RVYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVENTPKYAGVSPDRCKL 644
QY 619 ICRANTGTFFYLAPKVDGTLCSDDSTSVGVGKCIKAGCGNLGSKKFPDKCGVCGGD 678
Db 645 TCEAKGIGYFFVLQPKVVDGTFCSDDSTSVGVGQCVKAGCDRIIDSKKKFKDKCGVCGGN 704
QY 679 NKCKKVTGLFTKPMHGFNFVAIPAGASSIDIRGVYKGLIGDDNYLALKNQCKYLLN 738
Db 705 GSTCKKMSGIVTSRPGYHDVITIPAGATNIEVKHNRQGSERNQSFIAIRAADSTYILN 764
QY 739 GHEVVSVERDLVYKSLRYSGTGTAVESLQASRILEPLFVLSVGKMTTPPRVYSF 798
Db 765 GNFTLSTLEQDLYKTVLRYSGSSAALERIRSFPLKEPLTIVLMVGHALRPKIKETY 824
QY 799 YLPKEPREDKSHHPDRGPSVLHNSVLSLNQVEQDPDRPARVAGSWGSPCSACSGS 858
Db 825 FMKKKTES-----FNAIPTFS-----EWVIEWGECSKTCGSG 857
QY 859 LQRRADVRCGSAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSGRGFOR 915
Db 858 WQREVQCRDINGH---PASECAKEVKYPASTRPCADLPCPHQVGDWSPCSTCKGKYK 914
QY 916 RSLKCVGHGGRLLARQCNLHRRKQBP-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCDPLAKPKRHVIDFCTLTQC 950

RESULT 2
T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, N.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
C:Accession: T00355
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-837 <ISH>
A:Cross-references: EMBL:AB014588; NID:g3327189; PIDN:BAA31663.1; PID:g3327190
A:Experimental source: brain
C:Genetics:
A:Gene: KIAA0688
C:Superfamily: thrombospondin type 1 repeat homology
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 37.1%; Score 1913; DB 2; Length 837;
Best Local Similarity 46.3%; Pred. No. 2.3e-122;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

QY 1 MLLGLITLTAFAAGTGGSEPEREVVPIRLDPDINGRRYWRGPDSDGQGLIFQITAF 60
Db 37 LLLLLASLPSARLASPLPREIEIVPEKLGSLV-----PGSGTFAIRLLCRLQAF 88
QY 61 QEDFYHLITPDAQFLAPASTSTHGLGVPLQGLTGSSDRLRCFYSGDVNNAEPDSFAVSLC 120
```

```
Db 89 GETLLLEQDSGVQVBEGLTVQYLGQAPE--LLGGAEP--GTYLTGTINGDPESVASLHWD 145
QY 121 GG-LRGAFGYGAEYVTSPLPNASAPAAQRNSQAHLLQRRVPGGPGSGDPTSCRGVASG 179
Db 146 GGALLGVLYRGAELHLOPLEGGTPNSA--GGPCAHLIRK-----SPASOGGPMCNV--- 196
QY 180 WNPAILRALDPYKPRRAGFGESRRSRRGRKRFVTSIPRYVETLVVADESXVNFHGDLE 239
Db 197 -----KAPLGSPSPRPR--RAKRFASLSREVETLVVADDKMAAFHAGL 239
QY 240 HYLLTILATAARLYRHPSILNINIVVVKVLLLRDRDSGPKVTGNAALTURNECANOKKL 299
Db 240 RYLLTVMAAAKAFKHPISIRNPVSLVTVRLVILGSGEGPGVGSAAQTURSCFAWORG 299
QY 300 NKVSDKPEYWDTAILFTRODLGCATTCDTLGMADVGTMDCKPRKSCSVIEDDGLPSAFIT 359
Db 300 NTPEDSDPHDFDAILFTRODLGCVSTCDTLGMADVGTCDPARSCAIVEDDGLQSAFTA 359
QY 360 AHELGHVFNPHDNVNVCEEVFKL-RANHMSPTLIQIDRANPWSACSAAIITDLDG 418
Db 360 AHELGHVFNHLDNSKPCISLNGPLTSTRHMAPVMAHVDPPEEPWSPCSARFITDLDNG 419
QY 419 HGDCLLDQPSKPSLPEDLPGASVTLSSQCELAFGVSKPCPYM-OYCTKLWCTGKRAKQ 477
Db 420 YGHCLLDKPEAPLHLPVTFPGKYDADRCQLTFGPDSPRHCPLPPPCAALWCSGLN 479
QY 478 MVCOTRHPWADGTCGEGKCLKACVERHNLNKHRVD--GSWAKWDYGPCSRTC 535
Db 480 AMCQTKHSPWADGTCGPAQACGMGRCLHMDQLQDENIPQAGGWPWGWDGDCSRTCGG 539
QY 536 VOLARROCTNPTPANGKYCEGVYKYSRNLPCPSASGSKSFREBOCEAFNGYNHSTN 595
Db 540 VQFSRDCRTPVPNGGKYCEGRTFRSCNTEDCP-TGSALTFRREQCAA---YNHRTD 595
QY 596 RLTL---AVAWPKYSGVSPRDKCLICRANGTYFYVLAPKVVVDGCLSPDSTSV 652
Db 596 LFKSFGPMDWPRYTGAVPDQCKLTCQARALGYIVVLEPRVVDGTPCSPDSSVCVQ 655
QY 653 KCIKAGCDNGLGSKKRFDCGCGGDNKSKKVTGLFTKPMHGYNFVVAIPAGASSIDIR 712
Db 656 RCIHAGCDRIIGSKKFKDKCMVCGDGGSGSKSGSRFRKFRYGYNNVVTIPAGATHLVR 715
QY 713 QRGYKGLIGDDNYLALKNQSKYLLNGHFVVSAVERDLVVKGSL-LRYSCTGTAVESLQ 771
Db 716 QQGNPG--HRSIYLALKLDPGSGYALNGEYTLMPSPDTPVLPGAVSLRYSGATRASETLSG 773
QY 772 SRPILPDLTVELSVGKMTTPRVRYSFYLPKPREDKSSHPKD 814
Db 774 HGPLAQPLTLQVLVAGNPQDTRLAYSFFVRPTSTPRTPQD 816

RESULT 3
T47158
hypothetical protein DKFzp762C1110.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47158
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: 224379
A:Accession: T47158
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-550 <AAA>
A:Cross-references: EMBL:AL162080
A:Experimental source: adult melanoma (MeWo cell line); clone DKFzp762C1110
C:Genetics:
A:Note: DKFzp762C1110.1

Query Match 31.1%; Score 1604; DB 2; Length 550;
Best Local Similarity 49.7%; Pred. No. 1.6e-101;
Matches 284; Conservative 106; Mismatches 142; Indels 40; Gaps 11;
```


QY 934 NLHRKPOELDFCVLRPC 950
Db 1060 GREQKATERECNRIPC 1076

RESULT 5
T18517
procollagen N-endopeptidase (EC 3.4.24.14) I - bovine
N:Alternate names: procollagen N-proteinase
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18517
R:Collige, A.; Nusgens, B.V.; Lapiere, C.M.
submitted to the EMBL Data Library, February 1996
A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.
A:Reference number: Z18941
A:Accession: T18517
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1205 <COL>
A:Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253.1
A:Experimental source: skin
C:Genetics:
R:Gene: PC I-NP
C:Function:
A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to
C:Keywords: hydrolase; metalloproteinase

Query Match 23.1%; Score 1192; DB 2; Length 1205;
Best Local Similarity 31.7%; Pred. No. 4.8e-73;
Matches 329; Conservative 134; Mismatches 403; Indels 172; Gaps 42;

QY 1 MILLIGTLTAFAGTAG-----GSEPERVVVPIRLD----- 32
Db LLLLLLLPLDARLAAAADPPGPGHGCAERILAVPRDQAQGLVSHVVSAATAPAG 75

QY 33 -----PDINGRRYWRGPEDSGDGLIFQIFQAFEDFYHLTPDAQFLAPAFST 81
Db VTRPRAAPAQIPCLSG-----GSEEDPGGLFYNNVTFGRDLHLRLRPNARLVAPGATV 129

QY 82 EHLGVPLQGLTGSSDLRCFVSGDV-NAEPDSFAVSLCGGLRGAFYRGAEYVISP 140
Db 130 EWQGE--SGATRVPELLGTCLVGDVAGLAESSVALSNCDSLGLIRMEEEFFIEPLE 187

QY 141 NASAPAAORNSOG-AHLQRRGVP-----GGPSGDPTSRGAVSGWNPAILRALDPK 193
Db 188 KGL--AAKAEQGRVHVYHRTTSPPLGGPQALDT---GISADSLDSLRL----- 237

QY 194 RRAGFGESRRRSRRAKRFVSIPIRY-VETLVVADESVMKFGAD-LEHYLLTLTAAAR 251
Db 238 ---GVLEERVNSRRMRRAADDYDNYEVLLGVDDSVVQFHGTEHVOKYLLTLNIVNE 294

QY 252 LYRHPILAPINIVVVKVLLLRDRSGPKV-TGNAALILRNECANOKKLNKVSDDKHPEW 310
Db 295 IYHDESLGAINVILRIILLYSGKMSLIETINPSSQLENVCRWAYLOQKPDTHDETH 354

QY 311 DTAILFTRODLGATCTDILGMADVTGMDPKRSCSVIEDDGLPSAFITTAHELGHVFNMP 370
Db 355 DHAIFLTRQDF-GPSGMQ--GYAPVTGMCHPVRSCTLNHEDGFSSAFVVAHETGHVLMGE 411

QY 371 HDNW-KVCEVFQKLRANHMSPTLTQIDRANPWSACSAIITDFLDSHGDCLLDQP--- 427
Db 412 HDQGNRCRGD---EVRIGSLMAPLVQAAPFRHFRHWSRCSQOELSRYLHS---YDCLRDDEFT 466

QY 428 -SKPISIPEDLPQASVTLTQOCEAFRGVSKPCPYMQY---CTKLWCTGKAKGMVQOTR 483
Db 467 HDNP-ALPQ-LPGLHYSMEQCRDFGLGMMCTAPRTFDPCKQLMCS-HPNPYFCKTK 523

QY 484 HFPWADGTSGEGKCLKGACV-----ERHNLNKHRYDGSWAKWDYPGCSRTCGGV 536
Db 524 KGPPLDGTMCAPKCHCFKGHCITLTPDLIKR-----DGNWGAWSPPGCSRTCGTV 575

QY 537 QLARRQCTNPTPANGKYCEGVVVKYRSCNLBPCPSSASGKSFREQCEAFNGY----- 590
Db 576 KFRTRQCDNPHANGRTCSGLAYDFQLCNSQDCPDALA--DFREQCROWDLIFEHGA 633

QY 591 -NHSTNRLTLAVAWPKYSVSPDRCKKLCICRANGTGYFVVLAPKVVDDTCLCS-PDSTSV 648
Db 634 OHH-----WLP-HEHRDAKERCHLYCESKETGEVVSVMKRVHDTGRCYKDAFSL 682

QY 649 CVQGGKCIKAGCDNGLSKRFRDKCGVCGDNKSKCKVTGLFTKPMH--GYNFVVAIPAGA 706
Db 683 CVRGDCRKYCGDVIGSSKQEDKCGVCGDNHSHCKVVKGTFSRSPKGLGYIKMFEIPAGA 742

QY 707 SSIDIRQRYGKGLIGDNYLALANSQ-GYLLNGHFVVSVAERDLVVKGLSLKYSCTGTA 765
Db 743 RHLLIOEAD-----TSSHLLAVKNLETGRFILNEENDVPNSKFTFAMGVWEYRDE-DG 796

QY 766 VESLOASRPITLPTVELSVGKMTTPRVRYSPFLKPEKREDKSSHPKDPGPPSVLHNSV 825
Db 797 RETLQTMPLHGTTIVLVIPG---DARISLTY-----KYMIHEDS 834

QY 826 LSLSNQVEQPDPRPPARWAGSWGCSGSLQKRAVDCRGASAGQRTVPA--CDAAGR 883
Db 835 LNVDDNNVLEDDSVGVWEALKWSPCKSPCGGSGQFTKYGCRRLDHRKVMVHRGFCDSVSK 894

QY 884 P-----VETQACGECPTWELSAWSPCKSCGR-GFQRESLKV---CHGGRLLAFDQ 932
Db 895 PKAIRRTCNQECQ--PVWVTGEMEPSCSRSGRTGMQVSRVRCVOPLHNNTTTSRVHIXH 952

QY 933 CNLHRKPOELDFCVLRPC 950
Db 953 CN-DARPEGRACNRRLC 969

RESULT 6
C89114
protein C37C3 6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89114
R:anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustli.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: C89114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1558 <STO>
A:Cross-references: GB:chr_V; PIDN:AA25867.1; PID:g3294501; GSPDB:GN00023; CESP:C37C
C:Genetics:
A:Gene: C37C3.6a
A:Map position: 5

Query Match 13.2%; Score 681; DB 2; Length 1558;
Best Local Similarity 33.7%; Pred. No. 4.3e-38;
Matches 162; Conservative 63; Mismatches 176; Indels 80; Gaps 19;

QY 517 GSWAKWDPYPCSTRCGGVQOLARRQCTNPTPANGKYCEGVVVKYRSCNLBPCPSSASG 576
Db 77 GNGWPWPVPENCSSRSGGVQLEKRCQSGD-----CTGASVRYISCLNAC---ESG 125

QY 577 KSFREQCEAFN-----GYHNSTNRLTLAVAWPKYSVSPDRCKKLCICRANGTGYFVL 631
Db 126 TDFRAEQCSKFNDEALDGNVHK-----WTP-YKG---KNKCELVCPESGNFYKW 172

QY 632 APKVVDGTCLSPDSTSVCSVOGKCIKAGCDNGLSKRFRDKCGVCGDNKSKCKVTGLFTK 691
Db 173 ADKVVVGTCKDCKNDKNDICVDGECPLVCGDKGLSSLAFLDKCGDKGDSCTCKTIEGRFE 232

QY 692 P--HDGYNFVVAIPAGASSIDIRQRYGKGLIGDNYLALANSQGYLLNGHFVVSVAERD 749
Db 233 RNLSPGVHDIILKPEGATINIKQEAR-----KSTNNLALNGSDHFLYINGLNGLIQ-VEKE 286


```

Qy 792 PRVRYGYIPKPEPDKSHHPDGRGSPVLHNSVLUSNOVEDPCDRPP-----ARKWA 845
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 705 NELRITFF-----KARSYIMGSRVCRNKKRTITSEKNIEVRSCINGPCNAIGVWGTV-- 756
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 846 GSWGPCSCSGSG--LQKRAVD---CRGSAGORTVPACDAAHPPVETQACCEPCPT---W 897
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 757 GGMSTGCTSCGPGTLLVRQTCRNEPCDGSAHERR--SCNVA-----TCNDNGIW 803
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 898 EL-SANSPCKSCGGRGFORRLSKVCVGHG--GRLLARDQCN 934
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 804 SLWNWSDCSRVCGGLRSLRSCFCGSGCMGASSEQQFCN 843
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
T22545
hypothetical protein F53B6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22545
R;White, S.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19578
A;Accession: T22545
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1059 <WIL>
A;Cross-references: EMBL:Z81086; PIDN:CAB03121.1; GSPDB:GN00019; CESP:F53B6
A;Experimental source: clone F53B6

```

Db 230 SFAENRKTDLLETRKKH-----DNAQLLTAIDFNG-----PTIGYAVIGSMCHPKRSVGIQVQD 281
QY 351 ---DGLPSAFTTAHELGHVFNPHDNVNVKVEEVECKLRA NH---MMSPTLIQIDRANPWS 404
Db 282 YSPINLVAVVIAHMGHNLGHIHDT-----GSCSDGYPCIMGPT- ISNEPSKFFS 332
QY 405 ACSAAITDFLDSGHGDCLLDQP-----SKPISLPEDLPGASVTLSSQOCEL 458
Db 333 NCSYIQWDFIMNHNPECLINEPLGTDIISPPVCGNELL-----EVEECD- 378
QY 459 CPYQWYCTKLWCTCKAKQWVCQTRHPFPWADGTSCGEGKCLCKGACVERHNLAHRVD-- 516
Db 379 CGTPENCQNECCDA-----ATCKLK-----SGSQCGHGD-----CCEQCKFKSGRECR 422
QY 517 GSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGYGCVGRVVKYKSCNLEPCP----- 571
Db 423 ASMSECDPAEHCCTGQSS-----ECPADVFNHNGQPCLD-----NYGYCYNGNCPIMYHQ 471
QY 572 -----SSASGKSFREOCEAFNGYHNHSTNRLTLAVAVPKYSGVSPRD-KC-KLIC 620
Db 472 CYALFGADVIEAEDSCFKQKNGYGYCRKENGKKTIPCA-----PEDVKCGELYC 522
QY 621 RANGTG-----YFYVLAPK-----VVDGTLCSPDSTSVGVQKCI 655
Db 523 KDNPSGQNNPKMFYSNDEHKGMLPFTKCA--DGKVCVSNHGV 565

RESULT 14

S55270

C:Species: Crotales atrox - western diamondback rattlesnake

C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000

C:Accession: S55270

R:Zhou, Q.; Smith, J.B.; Grossman, M.H.

Biochem. J. 307, 411-417, 1995

A:Title: Molecular cloning and expression of catrocollastatin, a snake-venom protein from

A:Reference number: S55264; MUID:95251603; PMID:7733877

A:Accession: S55270

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-609 <2HO>

A:Cross-references: GB:U21003; NID:g710353; PIDN:AAC59672.1; PID:g710354

C:Superfamily: mouse meltrin alpha; disintegrin homology

Query Match 6.6%; Score 340.5; DB 2; Length 609;
Best Local Similarity 22.4%; Pred. No. 2.2e-15;
Matches 160; Conservative 100; Mismatches 282; Indels 17; Gaps 32;

QY 1 MLLGILTALFAGRTA-----GGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDOGLIF 55
Db 4 VLLVTICLAAPPYGGSSIIILESGNVNDYEVYIPKVTALPKG-----AVQPKYEDAMQY 57
QY 56 QITAFQEDFYLHLPDAQFLAPAFSTEHLGVPLQGTGGSSDLRRRCYSGDVNAEPDSPA 115
Db 58 ELKVGPEVVLHLGKNGKGLFSKDYSETHYSPDGREITTYPLVEDHCYHGRINDADSTA 117
QY 116 AVSLCGGLRGAFGYRGAEYVISP--LPNASAPAAQNSQGAHLQRRGVGPGSGDPTSR 173
Db 118 SISACNLGKHGFKLOGEMYLIEPLKLPDSEAHAVYKENV-----KEDIALKM 166
QY 174 CGVAGWNPAIRALDPYKPRRAGFESRRRSRAKRFVSIIP-----RYVETLVV 225
Db 167 CGVTQNW-----ESYEPIK-----KASQLVWTAEHQKYNPRFVELFLV 205
QY 226 ADESMVKFHGADLEHY---LLTLATAARLYRHPSTILNINIVVVKVLLLRDRDSGPKVT 282
Db 206 VDKAWVTKNNGDDJDKTRMYEIVNTVNTYRYM-----IHVALVGLIEWSNED---KIT 258
QY 283 --GNAALTILNFCAWOKKLNKVDKHPYWDTAIFLTRODLGATTCDTLGMADVGTMCD 340
Db 259 VKPEAGYTLNAGFENWRKTLTRKKH-----DNAQLLTAIDL-----DRVIGLAYVCSMCH 309

QY 341 PKRSCSVIED---DGLPSAFTTAHELGHVFNPHDN-VKVCBEVEFGKLRANHMMSPTLIQ 396
Db 310 PRKSTGIQDYSEINLVAVVIAHMGHNLGINHDSGYCSCGDYACIMRPEISPEFSTF- 368
QY 397 IDRANPWSACSAAITDFLDSGHGDCLLDQP-----SKPISLPEDLPGASVTLSSQOCEL 450
Db 369 -----FNSCYFECWDFIMNHNPECLINEPLGTDIISPPVCGNELL-----EVEECD- 416
QY 451 AFGVGSKPCPYQWYCTKLWCTCKAKQWVCQTRHPFPWADGTSCGEGKCLCKGACVERHNLA 510
Db 417 -----CGTPENCQNECCDA-----ATCKLK-----SGSQCGHGD-----CCEQCKKF 452
QY 511 NKRHVD--GSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGYGCVGRVVKYKSCNLE 568
Db 453 SKSGTECRASMSECDPAEHCCTGQSS-----ECPADVFNHNGQPCLD-----NYGYCYNG 501
QY 569 PCP-----SSASGKSFREOCEAFNGYHNHSTNRLTLAVAVPKYSGVSPRD 614
Db 502 NCPIMYHQCVDLFGADVIEAEDSCFERNQKNGYGYCRKENGKKTIPCA-----PED 552
QY 615 -KC-KLICRANGTG-----YFYVLAPK-----VVDGTLCSPDSTSVGVQKCI 655
Db 553 VKCGRLYCKDNPSGQNNPKMFYSNDEHKGMLPFTKCA--DGKVCVSNHGV 603

RESULT 15

S60257

C:Species: Mus musculus (house mouse)

C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000

C:Accession: S60257

R:Yagami-Hiromasa, T.; Sato, T.; Kurisaki, T.; Kamijo, K.; Nabeshima, Y.I.; Fujisawa-

Nature 377, 652-656, 1995

A:Title: A metalloprotease-disintegrin participating in myoblast fusion.

A:Reference number: S60257; MUID:96026308; PMID:7566181

A:Accession: S60257

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-903 <YAG>

A:Cross-references: EMBL:D50411; NID:gl054586; PIDN:BAA08912.1; PID:gl054587

C:Superfamily: mouse meltrin alpha; disintegrin homology

F.421-503/Domain: disintegrin homology <DIS>

F.349/Active site: Glu #status predicted

Query Match 6.3%; Score 324; DB 2; Length 903;
Best Local Similarity 21.5%; Pred. No. 4.6e-14;
Matches 159; Conservative 73; Mismatches 246; Indels 262; Gaps 27;

QY 2 LLLGILTALFAGRTA-----GGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDCG 52
Db 15 LLLALAGALLAPRARGSLWDQRAYEVARSL--LSKDPGIPGQSI-----PAKDHPDV 68
QY 53 LIFQITAFQEDFYLHLPDAQFLAPAFSTEHLGVPLQ-----GLTGGSSDLRRRCFYSGDV 107
Db 69 LTVQLQLESRLDILSLERNEGLIANGFETETHY---LQDGTDSLTRNHTD--HCYHGHV 123
QY 108 NAEPSFPAVSLCGGLRGAFGYRGAEYVISPPLNAS-----APA-AQRNSOGAHLQRRG 161
Db 124 OGDAAVSVLSLTCSDLRGLIMFENKTYSLPEPMKNITDSYKLVPAESMTNIQGL----- 176
QY 162 VPGPGSGDPTSRGCVASGNWPAIRALDPYKPRRAGFESRRRSRAKRFVSIIPRYVE 221
Db 177 -----CGSQHNKSNLTMDVSP-----GTSQMRARRHK---RETLKMTKVE 215
QY 222 TLVVADESMVKFHGADLEHYLLTLATAARLYRHPSTILNINIVVVKVLLLRDRDSGPKV 281
Db 216 LVIVADNREFQGRKDLKLVKQRLIEIANHVDKFRPLN-IRIVLVGVGVWNDIHK-CSI 273
QY 282 TGNRAALTILNFCAWOKKLNKVDKHPYWDTAIFLTRODLGATTCDTLGMADVGTACDP 341
Db 274 SODPTRLHEFDLWRRKIKLLPRKSH-----DNAQLISGYVFOGT---TIGMAPINSMCTA 325
QY 342 KRSCSVI---EDDGLPSAFTTAHELGHVFNPHDNV-KVCBEVEFGKLRANHMMSPTLIQ 397

```
Db 326 EOSGGVWMDHSDSPLGAATVLAHELGHNFNMHDTLERGCSORMAAEKGGCIMPNS----- 381
QY 398 DRANP-----WSACSAIITDFLDSHGDCLLDOP----- 427
Db 382 -TGFPFPMVFSSCRKDLFASLEKGMGMLFNLPEVKOAFGGRKCGNGYVEEGEECDGCE 440
QY 428 -----SKPISLPEDLPGAS--- 441
Db 441 PECTNRCCNATCTLKPDVCAHGOCCEDCOLKPPGTACRGSNSCDLPEFCTGTAPHC 500
QY 442 -----YTLQQCELAFGVGSKPCPYMOY----- 464
Db 501 PANVYLHDGHPGQGVGYNGICOTHEQQCVTLWGPAGKAPAGICFERVNSAGDPYGNC 560
QY 465 -----CTKLWCTGKAK-----COMVCQTRHEPW 487
Db 561 GKDSKSAFAKELRDAKCGKIQCGGASRPVIGTNAVSIETNIPQOEGGRILCRGTHVYL 620
QY 488 AD-----GTSCGEGKCLKGACYERHNLNKHRYDGSNAKWDPYCGSRTC-GGGV 536
Db 621 GDDMPDPGLVLAGTKCAEGKICLNRRQNISVFGVHK-----CAMQCHGRGV 667
QY 537 QLARROCTNPTPANGKYCE 556
Db 668 CNNRKNC-----HCE 677
```

Search completed: May 9, 2003, 15:24:10
Job time : 38 secs

GenCore version 5.1.1-p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:19:01 ; Search time 45 Seconds
(without alignments)
4349.884 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLTLTAFAGTAGGSE.....DOCNLRHPQELDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTRMBL.21.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviris.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5162	100.0	950	4 Q8TE58	Q8TE58 homo sapien
2	1685	32.6	340	11 Q91256	Q91256 mus musculu
3	1587	30.7	2165	5 Q19791	Q19791 caenorhabdi
4	1295.5	25.1	1081	4 Q8TE60	Q8TE60 homo sapien
5	1247.5	24.2	1072	4 Q8TE57	Q8TE57 homo sapien
6	1239	24.0	1688	5 Q8SXB0	Q8SXB0 drosophila
7	1220	23.6	1095	4 Q8TE56	Q8TE56 homo sapien
8	1206.5	23.4	1229	5 Q9VF61	Q9VF61 drosophila
9	1185.5	23.0	1054	5 Q9W493	Q9W493 drosophila
10	1165.5	22.6	1207	4 Q8TE59	Q8TE59 homo sapien
11	1122	21.7	1223	4 Q8WXS8	Q8WXS8 homo sapien
12	1120.5	21.7	1159	4 Q8TEY8	Q8TEY8 homo sapien
13	1118	21.7	1223	4 Q8TE55	Q8TE55 homo sapien
14	893.5	17.3	1427	6 Q96L37	Q96L37 homo sapien
15	830	16.1	269	6 Q9GL54	Q9GL54 oryctolagus
16	706	13.7	1280	11 Q9EPX2	Q9EPX2 mus musculu

17	693	13.4	1235	4	Q95428	Q95428 homo sapien
18	681	13.2	2167	5	Q76840	Q76840 cu-norhabdi
19	663.5	12.9	1572	5	Q44938	Q44938 laemonchus
20	599	11.6	1062	5	Q19204	Q19204 caenorhabdi
21	581.5	11.3	2174	5	Q9GQRO	Q9GQRO drosophila
22	579.5	11.2	3060	5	Q9AV4	Q9AV4 drosophila
23	569.5	11.0	766	4	P82987	P82987 homo sapien
24	559.5	10.8	1444	5	Q17591	Q17591 caenorhabdi
25	558	10.8	192	6	Q5N24	Q5N24 equus cabal
26	557	10.7	3198	5	Q9U8G8	Q9U8G8 manduca sex
27	550.5	10.7	525	4	Q96RW4	Q96RW4 homo sapien
28	518	10.0	187	6	Q95N23	Q95N23 equus cabal
29	511.5	9.9	1091	5	Q9W126	Q9W126 drosophila
30	510.5	9.9	872	5	Q22580	Q22580 caenorhabdi
31	493	9.6	951	4	Q60345	Q60345 homo sapien
32	487.5	9.4	1014	5	Q9SR33	Q9SR33 drosophila
33	426	8.3	790	5	Q8T458	Q8T458 drosophila
34	411.5	8.0	1059	5	P90884	P90884 caenorhabdi
35	410	7.9	183	11	Q99JP1	Q99JP1 mus musculu
36	402.5	7.8	364	4	Q9UG01	Q9UG01 homo sapien
37	379	7.3	899	13	Q8UVF1	Q8UVF1 coturnix co
38	365	7.1	922	13	Q8UVF2	Q8UVF2 coturnix co
39	356	6.9	610	13	Q93523	Q93523 bothrops ja
40	356	6.9	610	13	Q8QG88	Q8QG88 bothrops in
41	343	6.6	610	13	Q9YI20	Q9YI20 agkistrodon
42	343	6.6	622	4	Q9H8X0	Q9H8X0 homo sapien
43	340.5	6.6	609	13	Q90282	Q90282 croctalus at
44	340.5	6.6	609	13	Q9W6M5	Q9W6M5 agkistrodon
45	338	6.5	192	11	Q9CX59	Q9CX59 mus musculu

ALIGNMENTS

RESULT 1

Q8TE58 PRELIMINARY; PRT; 950 AA.

AC Q8TE58;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Metalloprotease disintegrin 15 with thrombospondin domains.

GN ADAMTS15.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21856482; PubMed=11867212;

RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,

RA Lopez-Otin C.;

RT "Cloning, expression analysis, and structural characterization of

seven novel human ADAMTS, a family of metalloproteinases with

disintegrin and thrombospondin-1 domains.";

RL Gene 283:49-62(2002).

DR EMBL; AJ315733; CAC86014.1; .

KW Integrin; Protease.

SQ SEQUENCE 950 AA; 103286 MW; 5DFBE18285CCCC3B CRC64;

Query Match 100.0%; Score 5162; DB 4; Length 950;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLTLTAFAGTAGGSEPEVVPVIRLDPDINGRRYRWGPDGSGQGLIFQITAF 60

Db 1 MLLGLTLTAFAGTAGGSEPEVVPVIRLDPDINGRRYRWGPDGSGQGLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTEHGVPLOGLTGGSSDLRRCFYSGDVNAEPDFAAVSLC 120

Db 61 QEDFYHLTPDAQFLAPAFSTEHGVPLOGLTGGSSDLRRCFYSGDVNAEPDFAAVSLC 120

QY 121 GGLRGAFYRGAEYVISPLPNASAPAAQRNSQGAHLQLRRGVPGPGSDPTSRGCVASGW 180

```
|||||
Db 121 GGLGAFGRCAEYVISPPLNAPASAPQNRSGAHLQLRRGVGPGSGDPTSRGCVASGW 180
Qy 181 NPAILRALDPYKPRRAGFSGESRRSRGKRAKRFVSIPIRYVETTLVVADESVMVKFHGADLEH 240
Db 181 NPAILRALDPYKPRRAGFSGESRRSRGKRAKRFVSIPIRYVETTLVVADESVMVKFHGADLEH 240
Qy 241 YLLTLATAARLYRHPSTLNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLN 300
Db 241 YLLTLATAARLYRHPSTLNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLN 300
Qy 301 KYSDKHPEYWDATILFTRODLGGATTCDTLGMADVGTMCDFKRSVCVIEDDGLPSAFTTA 360
Db 301 KYSDKHPEYWDATILFTRODLGGATTCDTLGMADVGTMCDFKRSVCVIEDDGLPSAFTTA 360
Qy 361 HELGHVFNPHNDNVKCEVEFGKLRANHMSPITLIQIDRANPWSACSAAITDFLDSGHG 420
Db 361 HELGHVFNPHNDNVKCEVEFGKLRANHMSPITLIQIDRANPWSACSAAITDFLDSGHG 420
Qy 421 DCLLDQPSKPTSLPDLPGASYTLSSQCELAFAFGVSGSKPCPYMQYCTKLWCTGKAKGQMV 480
Db 421 DCLLDQPSKPTSLPDLPGASYTLSSQCELAFAFGVSGSKPCPYMQYCTKLWCTGKAKGQMV 480
Qy 481 QTRHPFWADGTSGBGKLCCKGACVERHNLNKHVRVDSWAKWDYGPCSRCTCGGGVQJAR 540
Db 481 QTRHPFWADGTSGBGKLCCKGACVERHNLNKHVRVDSWAKWDYGPCSRCTCGGGVQJAR 540
Qy 541 RQCTNPTFANGGKYCEGVRYKYSRNLPCPSSASGSKSFREEQCEAFNGYNHSTNRLTLA 600
Db 541 RQCTNPTFANGGKYCEGVRYKYSRNLPCPSSASGSKSFREEQCEAFNGYNHSTNRLTLA 600
Qy 601 VAWVPKYSVSPDRCKLICKRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQKCIKAGCD 660
Db 601 VAWVPKYSVSPDRCKLICKRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQKCIKAGCD 660
Qy 720 GNLSGKKRFDKCGVGGDNKSKCKVTGLTTPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Db 720 GNLSGKKRFDKCGVGGDNKSKCKVTGLTTPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Qy 721 GDNTYALKNSGKYLNGHGFVSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLT 780
Db 721 GDNTYALKNSGKYLNGHGFVSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLT 780
Qy 781 VEVLVSGKMTPPRVRYSYFLPKPEPREDKSSHPKDPGRGPSVLHNSVLSLSNOVEQDDRRPP 840
Db 781 VEVLVSGKMTPPRVRYSYFLPKPEPREDKSSHPKDPGRGPSVLHNSVLSLSNOVEQDDRRPP 840
Qy 841 ARWVAGSWGPCASCSGSLQKRAVDCRSAGQRTVPACDAAHRPVTQACGEPCTWELS 900
Db 841 ARWVAGSWGPCASCSGSLQKRAVDCRSAGQRTVPACDAAHRPVTQACGEPCTWELS 900
Qy 901 ANSPCKSGCGRGFQRRLSKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950
Db 901 ANSPCKSGCGRGFQRRLSKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950

RESULT 2
Q91256 PRELIMINARY; PRT: 340 AA.
AC Q91256:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE Similar to a disintegrin and metalloproteinase with thrombospondin
DE motifs 1 (ADAMTS-1) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC009667.1; AAH09667.1; -.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 2.
KW Integrin.
FT NON_TER 1
SQ SEQUENCE 340 AA; 37021 MW; 06562C747634A8BD CRC64;

Query Match 32.6%; Score 1685; DB 11; Length 340;
Best Local Similarity 91.7%; Pred. No. 3.4e-133;
Matches 308; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

Qy 615 KKKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQKCIKAGCDGNLSGKKRFDKCGV 674
Db 5 KKKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQKCIKAGCDGNLSGKKRFDKCGV 64
Qy 675 CGGDNKSCKKVTGLTTPMHGYNFVVAIPAGASSIDIRQGYKGLIGDNYLALNSGK 734
Db 65 CGGDNKSCKARVGLTTPMHGYNFVVAIPAGASSIDIRQGYKGLIGDNYLALNSGK 124
Qy 735 YLLNGHFVYSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLTVEVLVSGKMTPPRV 794
Db 125 YLLNGHFVYSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLTVEVLVSGKMTPPRV 184
Qy 795 RYSFYLPKPEPREDKSSHPKDPGRGPSVLHNSVLSLSNOVEQDDRRPARWVAGSWGPCAS 854
Db 185 RYSFYLPKPEPREDKSTRPKDPRGSLVLRNSVLSLSNOVEQDDRRPARWVAGSWGPCSVS 244
Qy 855 CGSLQKRAVDCRSAGQRTVPACDAAHRPVTQACGEPCTWELSAWSPCKSGCGRGFQ 914
Db 245 CGSLQKRAVDCRDSFGQOGASACDVDRPLEKRAAGGEPCTWELGNWSPCKSGCGRGFK 304
Qy 915 RRLSKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950
Db 305 RRLPKCVGHGGRLLARDQCDLRRKPQELDFCVLRPC 340

RESULT 3
Q19791 PRELIMINARY; PRT: 2165 AA.
AC Q19791; Q27524;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadaty S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lighning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smailson N., Smith A., Sonhammer E., Staden K., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL; Z69361; CAA93288.1;
DR EMBL; Z69360; CAA93288.1; JOINED.
DR EMBL; Z69360; CAA93287.1;
DR EMBL; Z69361; CAA93287.1; JOINED.
```

```

DR HSP: P15167; IDTH.
DR MEROPS; M12.135; -.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPL.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 14.
DR SMART; SM00209; TSPL; 18.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSPL; 6.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN1.
DR SO SEQUENCE 2165 AA; 244397 MW; FCC3DA8AAA9C4888 CRC64;

Query Match 30.7%; Score 1587; DB 5; Length 2165;
Best Local Similarity 34.6%; Pred. No. 9.5e-124;
Matches 380; Conservative 149; Mismatches 382; Indels 186; Gaps 36;

QY 2 LLLIGILTAFATAGG--SPEREV-----VPI-----RLDPDIN-----GRR 39
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 18 LILLVCLVYALQSGGTISFSDSLFRSAKYSGVPHHSRWRODAGIHVIDSHHIVRR 77
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 40 --YYWRGPD-----SGDGLIFQITAFQED--FYLHL-----TPDAQFLA- 76
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 78 DSYGRGRKRDVTDSTDRRRRLQGVARDCGHACHRLRSDDAVYIVLHWNQIPIUSHKSV 137
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 77 PAFSTEHLGVPL-----QGLTGGSDL--RRCFYSGDVNAEPDSFAVSLC-----GGLRG 125
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 138 PHFSNSNF--APMVLVLDSEEEVRGMSRTDPCITYRAHVKG--VHOHSIVNLCDSGDGLYG 195
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 126 AFGYRGAEYVISPUNASA---PAAQRNSQAHLLORRYPVGGSGDPTSHCV----- 176
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 196 MLALPSGIHVPEIISNGTEHDGASHRQ--HLVRKFPDMHFKSFDHLNSTSYNETETT 253
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 177 ----ASGNPAILRALDPYKPRRAGFGEISRRSRGAKREFVSPRVETILVADESVMK 232
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 254 VATWQDQWEDVIER-----KARSRAAN-----SWDHVEVLVADTKMYE 294
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 233 FHGADLEHYLLTLTAARLYRHPSIILNPINIVVVKVLLLRDRDSQPKVGTGNAALTIRNF 292
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 295 YHGSLSDYVLTUFTVASIYRHOSLRASINVVVKVILIVKTENAGRPRIITQNAOQTLDQF 354
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 293 CAWQKLNKYSDKHPEYWDTAIFTRODLGCAT--TCDTGLGMADVGTMC DPRKSCSVIEDD 351
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 355 CRWQYYNDPDDSSVQHHVAILLTRKDICRSQKCDTLGLAELGTWCDMKSCAILEDN 414
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 352 GLPSAFTAHELGHVFNPHDNVKKVCEV-----FGKLRAN---HMMSPTLI 395
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 415 GLSAAFTIAHELGHVFSIPHDHDKCKCTYMPVANKVCKFQSTKDKTQFQNFHIMATPLE 474
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 396 QIDRANPWSACSAIITFDLDSHG--DCILLDQSPKISLPE----DLPGASYTL SQOCE 449
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 475 YNTHPSWSFCSAGMLERFLFNRRGQTQCLDQFVERRYEDVFVRDEPKKYDAHQOCK 534
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 450 LAFGVGSKPCPYMQYCTKLWCTGKAGQMVQCOTRHFHPWADGTSGECK--LCLKGACVER 507
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 535 FVEGPASELCYPMTCKRLWCATFYGSQMSCRTOHPWADGTPCDESRSMFCHGACVRL 594
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 508 HNLNKHRYDGSWAKWDPYGPCSRTCGGVQJARQCTNPTPANGKYCEGVRYKYSNCL 567
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 595 APESLTKIDQWGDWRSWSGCSRTCGGVQKGLRDGCDSPKPRNGGKYCVQQRERYSCNT 654
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 568 EPCPSSASGSKSFREOCEAFNGYN-----HSTNRLTLVAVAPKYSGVSPDKCKLICR 621
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 655 QECPWD--QPYREVOCSEFNKDDIGIQVASTN-----THWVPKYANVAPNERCKLYCR 707
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 622 ANGTGYFYVLAPKVDGTLCSPDSTSVCVQGKC IKAGCDGNLGSKRFRDKCGVCGGDNKS 681
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 708 LSGSAFYLLRKRVVDGTPCDRNGDDICVAGACMPAGCDQLHSLTLRROKCGVCGGDSS 767
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 682 CKVKTGLFTKP--MHGYNFVVAIPAGASSDIDIRGTYKGLIGDNNYLAKNSQKYLNGH 740
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 768 CKVKGTFNEQGTGYNEVYMKIPAGSANDIDROKYNMKEDDNYLSLRAANGFEFLNGH 827
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

```


Db 999 SHTCGKWRKRAVACKSTNPSARAQLLPDAVCTSEPKPRMHEACLQRC 1047

RESULT 6

Q8SXB0 PRELIMINARY; PRT; 1688 AA.

AC Q8SXB0; DB 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DE GH16393p.
OS CG6107.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuncio J., Pacieb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY094716; AAM11069.1;
SQ SEQUENCE 1688 AA; 189867 MW; 48FB8DD4DE0CA4D2 CRC64;

Query Match 24.0%; Score 1239; DB 5; Length 1688;
Best Local Similarity 32.0%; Pred. No. 1.2e-94;
Matches 321; Conservative 142; Mismatches 339; Indels 200; Gaps 39;

QY 55 FOITAFQEDFVHLTPDAQFLAPAFSTHGLVPLQG-----LTGSSDLRR--CFYSGD 106
DB 296 YNLNVFGQLHLVLRQDAFV--HNHSMTHRIKLEGEHPGTEAEAEORHLGCFYSGY 354
QY 107 VNAEPDSFAAVSLCGGLRGAFGYRGAEVISPLPNASAP-----AAORNSOGAHLQ 158
DB 355 VEDDPHSMVSVSLCGGTGIKTSFGALLIOPVNRSTSSDEVLRHVPKRSORNA--HVS 412
QY 159 RRGVPGSGDPTSRGCVASGNWPAILRALDPYKPRAGFGE-----SRSRRRSGRAKRFV 214
DB 413 KFEI-----GLDDFMKSLQVQEEQKSKRKLNRKRHYA 448
QY 215 SIPRYV---ETLVVADESVMKFGADLEHVLTLTAAALYRHPSTLNPINIVVVKVLL 271
DB 449 DVDNQVTVLELIANDSMKQFHGEDLPVILILMSIVSIFADASIGNSIRILLVRLIS 508
QY 272 LRDRDSGPV---TGNAAITLRNFCAMQKKLVSKHPYWDATLFTQDILCGAT--- 325
DB 509 L-----PNINDOTHSSNEMLKHFQCP---INQSYER-----DTAMLTITREPICGVPKG 555
QY 326 TCDPLGMADVGTCDPKRSCSVIEDGLPSAFTAHGLHGVFMHNDNVKVCVEVFGKLR 385
DB 556 ICHMGLGAEIGTVCS--SSSCSIVQDTGLPTAFTMAHELGHILNMHDDDKCMPPYVTRON 614
QY 386 AN---HMMSPTLIQIDRANPWSACSAAIITDLDSDHGDCLLDQPSKPSLIP---EDLPG 439
DB 615 NNKVLHIMSSVMGTHMHPWSKSRHFVSEFLEKTDKSL---ETSVGAHIPVGTRELPG 672
QY 440 ASYTLQOCELARGVSGKPCPYQYCTKLWC---TGKAKQMVCOIRHFFWADGTSCG-E 495
DB 673 EYISLDAQOCLSGFNDGFCPTDECKRLWCNRTSGNSNEQ---CASSNLEWADGTFCGSS 730
QY 496 GKCLKGACVERHNLNK-----RVDGSWAKWDYGFCSRTCGGGVGLARQCTNPTPANG 551
DB 731 GHWCQKQCVS---NKHGTVGQVNGWGWPTFTFCSLTTCGGGVGSRRECNOQVPENG 786
QY 552 GKYGCVRVYKRNLEPCPSSASGSKSFREOCEAFNGYHNSTNRLTLAVANVPKYSVGS 611
DB 787 GKYCTGSRKKYRCNTHQCPGSGMDP---REOQCYAMNGRMNIPGVPNTDKWPKYE--- 841

QY 612 PRDKCKLICRANGTGYFYVLAPKVVDTGLCSPDSTSVCVQKCIKAGCDNLSKKRFDK 671
DB 842 -KDACKLFCRMDMKVTFYMLKSMVTDGTSCAVDSFDKCVNGICRPAGCDNELNSIAKLDK 900
QY 672 CGVCGGDNKSKKVTG-LFTKPMHGYN-----FVVAIPAGASSIDIRGKGLI 720
DB 901 CGVCEGRNDTCHEVTGNLLVNSLLGLNDGNEPNKTIYYVTRIPKGSNIITQGYGYP--- 957
QY 721 GDDNYLALKNSQKYLNGHFVVSVERDLVVGSLRLYSVGTGTAVESLQA---SRPILP 778
DB 958 -DQNFIVLTDNRDNEELNGKF-LKTYPLKFVYAGVTQMOTGSSSVVEQVNTIYSWLSRD 1015
QY 779 LTVFVLSVGMKTPPR-----VRSFYLPKPRE- 806
DB 1016 LIVQIIISLDVSPSKRQDTVLLSYTYIDKPPDYAEVEIYRWEMQAPSNCDSLCEGRSHR 1075
QY 807 -----DKSSHK-DPRGPSV---LHNSVLSLSNQVQPDPPPARVW 844
DB 1076 LPACISTTQGVKVAPOFCDKSAMPKIDDRACNTDCLRLUTVTSISE----- 1121
QY 845 AGSWGPCSCSGS-GLQKRAVDC--RGSAGQRT---VPACDAHRPVEVTOACGEPCTW 897
DB 1122 -----CSAACGELGTREKTYACVQFTTNMORSNIIVDMYSYCKLKFVAYHEECREGC--W 1173
QY 898 ELSAWSPCSKSCGRGFORSLKCVGHGGLLARDQCNLHRKP 939
DB 1174 VLSEWSTCSKSCGTGSGQREAHCYLHNSR-VSDDLGNPRTKP 1214

RESULT 7

Q8TE56

ID Q8TE56 PRELIMINARY; PRT; 1095 AA.

AC Q8TE56; DB 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Metalloprotease disintegrin 17, with thrombospondin domains.
GN ADAMTS17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Odaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains";
RL Gene 283:49-62(2002).
DR EMBL; AJ315735; CAC86016.1; -.
KW Integrin; Protease.

SQ SEQUENCE 1095 AA; 121099 MW; A5C6540484754D5D CRC64;

Query Match

Best Local Similarity 23.6%; Score 1220; DB 4; Length 1095;

Matches 331; Conservative 139; Mismatches 386; Indels 190; Gaps 40;

QY 2 LLLGLTILAF-----AGRTAGGSEPEREVVPIRLDPDI-----NGRRYYWRGPE 46
DB 10 LVLPVLLLVWGLDGTAVGDAADVEVLPWRVPDVLPLPAAPCPRRRRRPTTP 69
QY 47 DS-----GQGLIFQITAFQEDFYHLTPDQFLAPAFSTHGLVPLGGLTGGSDLRRC 101
DB 70 AAPRAPGERALLHLPAFGRLDYQLRLRDLRFLSRGFEVEEAGAARR--RGRPAEL--C 125
QY 102 FYSGVDPNAPEDSFAVSLCGGLRGAFG--YRCAEVV-ISPPLNASAPAAQNSQAHLLQ 158
DB 126 FYSGVRLHGPGLSVLSACGAAGGLVGLIQGLQEQGVLOPLNNSQGFSGRE---HLIR 181
QY 159 RR-GVPGGFGSD---PTSRGCVASGNWPAILRALDPYKPRAGFGE-SRSRRRSGRAKRF 213

Db 182 RWSLTSPSAEAORPEQLCKVLT-----EKKPTWGRPSRDWRERRAIRL 228
Qy 214 VSPRYVETLVVAVESWVKFGAD-LEHYLLTLATARLYRHPISILNPINIVVKKVLL 272
Db 229 TS-EHTVETLVVADAMVQYGAEEAARFILTVMNVMYMFQHOSLGKIKINQIVTKVLL 287
Qy 273 RDRSGPKVTGNAALTLNFCAWOKK-----LNKV--SDKHPEYWDTAIFLTRQDL 321
Db 288 RORPAKLSIGHGERSLESFCHWNEEYGGARYLGNQVPGCKDDPLVDAAEVETRTDF 347
Qy 322 C-GATTCDTLGMADVGMCDPCKSCSVIEDGLPSAFTTAAHELGHVFNPHDNVKNVCEE 379
Db 348 CVHKDEPCDTGVIAVLGGVCSAKRCKVLAEDNGLNLAFTIAHELGNLGMNHD-----DD 402
Qy 380 VFGKLRANMSPTLIQIDRANP-----WSACSAALITDLDLSGHGDCLL---DQSPKPS 432
Db 403 HSSACGRSHIMSGEYV--GRNPSDLSWSSCRDLENFLKSKYSTCLLVTDPRSQTVR 460
Qy 433 LPEDLPGASYTLSCQELAFGVGSKPCPYMAY--CTKLWCTGKAGQWCVQTRFPWADG 490
Db 461 LPHKLPGMHYSANEOQILFGMNATFCRNMEHLMCAGLWCL--VEGDFSCCKTLDPLDGG 518
Qy 491 TSCGEGKLCGLKAGCVERNLNKHVRVDGSKWAKWDVPGPCSTCGGVQVLARQCTNPTPAN 550
Db 519 TEGGADKWCRCAGECVSKTPIEBH-VGDWSPWGAWSMCSRTCGTGARPRQRKCDNPPPGP 577
Qy 551 GKGYCEGVYKVRSCNLPSCSSASGSKFREDOCEAFNGYHNSTNRLTLAVAWPKYSGV 610
Db 578 GTHCPGASVEHAGVENUPCPKGL--PSFRQOQOAHDRLPKPKKGLTAV-----V 627
Qy 611 SPRDCKKLICRANGTYGYFVLAPKVVDTGLCSPDSTSVCGVGKCIKAGCDNGLSKKRF 670
Db 628 VDDKPELYCSPLGKESPLLVADRVLDTGCPGYETDLCVHGKQKICGCDGLIGSAAKED 687
Qy 671 KGVCGGDNKKCKVTGLFTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNDYLAKN 730
Db 688 RCGVCSGDGKTCHLVKGFDSH-----ARGTALKDSG-KGSI-----N 723
Qy 731 SQGKYLNGHFVVSVERDLVVKGLSILRGYSGTAVESLOSARPLEPLTVEVL-----S 785
Db 724 SDWKIELPEFOJA-----GTVRVKVRG-LWEKISAKGPKLPEHLMVLLFHDQD 773
Qy 786 VGMTPPRVRYSYFLPKPEKRDKSHKDPDR-----GPSV-----820
Db 774 YG-----IHYEYTPVNTAENQSEPKQDLSLFTWTHSGWEGCSVQCGGGERRTIVSC 827
Qy 821 --LHNSVLSLNOVEQDPRPP-----ARWAGSWGSCSASCGLQKRAVD 865
Db 828 TRIVNKTTLVNDSDCPQASRPEPOVRRCNLHPCOSRWVAGSPWSPCSATCEKGFQHREVT 887
Qy 866 C----RGSAGORTVPACDAARHPVETQAC-GEPC-PTWELSAWSPCSKSCRGFORSLIK 919
Db 888 CVYQLONGTHVATRPILYCGPRPAVQSCGQDCLLSINEASEWSQCSASCCKGVKRWIVA 947
Qy 920 CVHGGKRLIARDQCNLHRKPEQLDFC 945
Db 948 CTNSQCK-----CDASTRPRAEEAC 967

RESULT 8
Q9VF61 PRELIMINARY; PRT: 1229 AA.
ID Q9VF61;
AC Q9VF61;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE C66107 protein.
GN C66107.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.E.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer J.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Abayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cowley S., Dahlke C., Davenport L.B., Davies F.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.N., Woodage T., Worley K.C., Wu D., Yang S., Yac O.A.,
Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003709; AAF55199.1;
DR FlyBase; FBgn0038340; CG6107.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPl.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 2.
DR SMART; SM00209; TSPl; 3.
DR PROSITE; PS00215; ADAM_MEPRO; 2.
DR PROSITE; PS50092; TSPl; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc.
SQ SEQUENCE 1229 AA; 138136 MW; C1293180AB19B700 CRC64;

Query Match 23.4%; Score 1206.5; DB 5; Length 1229;
Best Local Similarity 32.0%; Pred. No. 4e-92;
Matches 321; Conservative 140; Mismatches 329; Indels 213; Gaps 43;

Qy 55 FOITAFQEDFYHLHTPDAQFLAPAFSTHEHGVPLQG-----LTGSSDLRR--CFYSGD 106
Db 195 YNLNVFGRQLHLVLRQDASFV-HNHSMTHIRILKEGEEHPGPEAEARHLCGFSY 253
Qy 107 VNAEPDGFAAVSLCGGLRGAFYRGAEYVISPPLNASP-----AAQNSOGAHLQ 158
Db 254 VEDDPHSMVSVSLCGGMTGIKTSFGALLIQPVNRTSSDVELHVRFKSQNRAR--HVS 311
Qy 159 RRGVPGSGDPTSRGCVASGNPAILRALDPYKPRAGFCE-----SRSRRRSRKRFV 214

Db 312 KPFL-----GLDDFMSKLEQVQEEQCKSKSRNLNKKRHYA 347
QY 215 SIPRYV---ETLVADSESVKFGHADLEHYLLTLTATARLYRHPSTILNPINIVVVKVLL 271
Db 348 DVDNQVYILEVLIANDSMKQFHGEDLOPYLILKMSIVSSIFADASGNSIRILLVRLIS 407
QY 272 LNDROSGPKV---TGNAAATLIRNCAMOKKLNKVDKHPYWDRAITFTRDOLCGAT--- 325
Db 408 L-----PNIINDQTHSSNMLKHFQCF---INQSYER---DTAMLITREPICGSPGK 454
QY 326 TCDTLGMDVQCTMDPKRSCSVIEDGLPLSAFTTAHELGVFNPNPHDNVKNVCEVFKLRL 385
Db 455 ICHMLGLAELGTVC--SSCSIVQDTGLPTAFTMAHELGLILNHNHDDDDKCMPIVTRQN 513
QY 386 AN---HMSPTL-IQIDRANPWSACSAIITDFLDSGHGDCULLDQPSKPSLSP---EDLP 438
Db 514 NKKVLHINSMVGIHM---HPWSKK---TD-----KCL--ETVGAHIPVGTTERLP 559
QY 439 GASVTLQOCELATGVGSKPCPYQYCTKLMC-----TGKAGQMVQCTRHFPWADGTSCG- 494
Db 560 GEIYSLDAQOCLSGFNGDFGYCTDECKRLWCNRTSGNSNEQ---CASSNLPWADGTPCGS 617
QY 495 EGKCLKGACVERHNLNKH---RVDSWAKWDYPGCSZTCGGVQVLARQCTNPTPAN 550
Db 618 SGHWCQKCKVCS---NKHGVRQVNGWGWPTEFTPCSLTCGGGVQVESRECCNQPVEN 673
QY 551 GKVCBGRVVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSV 610
Db 674 GKCYTGRKRYRSCNTHOCPPGMDP--REQCYAMNRMNMTPGVNPDTKWPKYE-- 729
QY 611 SPROCKLICRANGTGYFVVLAPKVDGTLCSPDSTSVCGKICAKCGNLSGKKRFD 670
Db 730 --KDACKLFCRMDKMTVFMKLSMVTGDTSCAVDSFDKCVNGICRPAGCDNELNSIAKLD 787
QY 671 KCGVCGDNKCKKVTG-LFTKPMHYN-----FVVAIPAGASSIDIRQGYKGL 719
Db 788 KCGVCEGRNDCHEVTGNLLSNLGLNDGNEPKNTIYVTRPKGASNLIIITQRTGP-- 845
QY 720 IGDNYIALKNSQKYLNGHFVSAVERDLVKGSLRLYSGRTATVESLQA--SRPIL 777
Db 846 --DQNFVILTDNRNELLNGKF-LKTYPLKFVYAGVTMQTSSSVVEQVNTIYSKLSR 902
QY 778 PLTVEVLSVGMKTPR-----VRYSFYLPKREPE----- 806
Db 903 DLIWOIISLDSVPSKRODTVLLSYSTYIDKPPDYAEVEIYRWEMQAPSNCDSLCEGRSH 962
QY 807 -----DKSSHPK-DPRGPSV---LHNSVLSLSNQVEQPDPRPARW 843
Db 963 RLPACISTGTGVKVPQCDKSAHPKIDDRACNTDCLNLTIVTSIEL----- 1009
QY 844 VAGSWGFCSCSGS-GLQKRAVDC--RGSAGQRT---VPACDAARHPVETQACGEPCT 896
Db 1010 -----CSAAGELGTREKTYACVQFTTNQRSNIVDMSYCKLKFQVAYHEBCEGCG-- 1060
QY 897 WELSAWSPKSCSGRQFQRLSKVGHGGRLLARDQCNLHRKP 939
Db 1061 WVLSEWTSKSCGTGSGQREAHCYLHNSR-VSDDLCLNPRTKP 1102
RESULT 9
Q9W493 ID Q9W493 PRELIMINARY; PRT: 1054 AA.
AC Q9W493;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE CG4096 protein.
GN CG4096.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RX Adams M.H., Celniker S.E., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Ananadis P.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Allrich J.F., Agayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Bailaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloeck A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harits N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchur K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AF003435; AAF46065.1; .
DR HSSP; P15167; IATL.
DR MEROPS; M12.231; .
DR FlyBase; FBgn0029791; CG4096.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 2.
DR SMART; SM00209; TSPI; 2.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSPI; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 1054 AA; 118616 MW; DC15455555CB6212 CRC64;

Query Match 23.0%; Score 1185.5; DB 5; Length 1054;
Best Local Similarity 30.5%; Pred. NO. 1.8e-90;
Matches 336; Conservative 140; Mismatches 375; Indels 249; Gaps 45;

QY 17 GGSEP-----EREVVVPRLDPP-----INGRYYWR---GPED 47
Db 23 GGVRLYGLHSELVAGGQLVPRVHPDGFMTWQLEYAHELDHRRHRRSLSNSCHD 82
QY 48 SGDOGLIFQITAFQEDFYHLHTPDQAFLAPAFSTFHLGVLQGLTGGSSDLRKFYSGEV 107
Db 83 TQAAIDLHLPLANETLHLEMAHSYFLAPNLVVERHRRDLRTSRPLTRHLNCHFGKV 142
QY 108 NAEPSFAVSLCGGL-----RGAFYRGAEYVISPL-----PNASA 144

Db 143 RGQATNVAISTCAGLVSCPNFCVLPELITLQVGHRTAGNEYFIEPSKEHEHPVNGH 202
Qy 145 P--AAORNS-OGAHLILOG--VPGGPGCD--PTSRCGVASGNPAILRALDPYKPRR-- 195
Db 203 PHVFORSSVKPHSLRKRKRGRGSGAEVSCGTR-- 248
Qy 196 -----AGFGES 201
Db 249 ETRLEWQARKVKVQGGROI RRRHHHHHHHHKHYRHHQOKISRVPHTKFKYEQFQTE 308
Qy 202 RSRRRSGRAKRVISIPRYVETLWVADESVMKFGADLEHYLTALTAARLYRHPHSILNP 261
Db 309 PDHAETPRRRSSISPRHETLIVADATMSAFH-RDLNGYLLTMMVSAKYKDSIGNS 367
Qy 262 INIUVKVLRLDRDGPX--VTGNAALTLRNFCAQMKLNKVSQDKHPYDWTAILFTRO 319
Db 368 IEIUVRIQLEDEESQQLNLTQNAOKNLDRCFSQWHLNKGSEKDPHHHDAVAILTRK 427
Qy 320 DLGATTCDTGLMADYGTMCPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDVKV-CE 378
Db 428 NIC-ANNMCTLGLANVGMCKPKOSCVNEDNGIMLSHTITHELGHNFMGFHD TAKIGH 486
Qy 379 EVFGKLRANHMGP TL-----IQIDRANPWSACSAAIITDFDSHGDCDLLDQSPKI--- 431
Db 487 PRVGP I--VHIMPTTFGADTLQV-----CWSNCSRKYITHELQGLGEC-LDDPPTPLDEY 539
Qy 432 SLPELDPGASYTLISOCELAFG-----VGSKPCPYMOYCTKLWCTGKAKGMQVQTRHF 485
Db 540 NYTGELPGMYNARGCRLQFNLTIDSEVGACSAFH-EFCSTLWC--KVNGE--CVTHNR 594
Qy 486 PWADGTCGEGKLCGLKAGCVERINLKNHRVDGSAKWDPYGPSCRTCGGGVQLARQCTN 545
Db 595 PTAPGTLGGRNKCQNGKCVRRREL--AAVNGGWGDWSESECSRGCGGVSTQORECDN 652
Qy 546 PTPANGKCEGVYRVKRSNLEPCSSASGSKSFEEQCEAFN--GYNSTNRLTLAVAW 603
Db 653 PVFANGGVFCIGERKTKICRKRCP--AEPSFRAQCCARFNVSYQATYK-----W 704
Qy 604 VPKYSGVSPRDKCLICRANGTGYFYVLAPKVYDGTLCSPDSTSVQVQCKTKAGCDGNL 663
Db 705 LPFEDKNP--CKLFCSDVDDTIANNWATVLDGTCTLTNNMCDIGTCKKVGCDWIV 761
Qy 664 GSKRRFDKCGVGGDNKSKKVTGLTKPM-----HGYNFVVAIPAGASSIDIRQYKGL 719
Db 762 DSEVQDDRCGVGSGSDQCPVRETYTDPFAAKDGAAYVEITVTPARAHILINE-----L 816
Qy 720 IGDNYLALKNSG--KYLNGHFWYSAVERDLWKGSLRLYSGTGTAVESLQASRPIL 777
Db 817 ANSPHFLAIATGGDRFYNGSLIS-MPGFEIAGAESLYDRVDEQ-ETITPQIQH 874
Qy 778 PLAVEVLSVGKMTPPRVRYSYFLPKPEPRDKSHPKDPRGPSVLHNSVLSLNSQVQPD 837
Db 875 SISLYAIYRGNESNAGIFYEFTLP-----ALN--VTAGRQFO--- 909
Qy 838 RPPARWAGSWGPCSSASGSGLOKRAVDC-----RGSAGO-----RTVP 876
Db 910 -----WLSNWTACSGGCVQHREPCQENKSGESNEFPQIVSKRNLLITAGTLP 964
Qy 877 ACDA--HRPV-ETQAGC-EPCPT-WELSAWSPCSKSGR-GF-----QRRSLKCVGHGG 925
Db 965 CWTAKNKRPARQSGCGDQCPAHWPWPQVFCPTVCRPVGFVAPQRRRSVVCLEHD 1024
Qy 926 RLLARDQCNLHRKQPELDFC 945
Db 1025 VVADAECGHLQKPAEMEPC 1044

RESULT 10
Q8TE59
ID Q8TE59 PRELIMINARY; PRT; 1207 AA.
AC Q8TE59;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ADAMTS-19.
GN ADAMTS19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTS, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
DR EMBL: AJ311904; CAC84565.1; -;
SQ SEQUENCE 1207 AA; 134061 MW; AP36F6BF586FDE2 CRC64;
Query Match 22.6%; Score 1165.5; DB 4; Length 1207;
Best Local Similarity 30.0%; Pred. No. 1.1e-88;
Matches 326; Conservative 145; Mismatches 380; Indels 235; Gaps 44;
Qy 16 AGGSEPEREVVPIRLDPDINGRR----- 39
Db 76 AAGSREVSVPVPLEEPVEGRSESLRPPPPSEGEDEBELESQELPRSGSAAALSPG 135
Qy 40 --YWRGSP-----EDSDQGLIFQITAEQEDFYHLHTPDQAF LAPAFSTE 82
Db 136 APASWQPPPPPPPPPPPAQHAEPDGDDE-VLLRIPAFSRDLYLLRRDRGRFLAPRAVE 194
Qy 83 HLGVPLOGLTGGSDLR-----RCFSGDVNAEPDSFAVSLCGGLGAGFYRGAEYV 135
Db 195 QRPNPGPPTGAASAPAPPADGCFYTGAVLRHPSLASFTCCGGGLMGFIOLNDEFI 254
Qy 136 --ISLPNASAPAA-----QRNSGAHLQRRGVPGSGDPTSRGCVASGMNPAILRA 187
Db 255 FIEPLNDTMAITGHPRVYRKRSMEEKVTEK-----SALHSHYCGIIS----- 298
Qy 188 LPDYKPRRAGFGESESRSSRRGRKRF-VSIPR--YVETLVVADESVMKFGAD-LEHYLL 243
Db 299 -DKGRP-----SRKIAESGRGRYSYKLPQENIETVVVADVAMPVSYHGADAAARFIL 351
Qy 244 TLLATAARLYRHPSTILNPINIVVVKVLLLRDRDGPVKVTGNAALTLLRNFCAWQ-KKLNKV 302
Db 352 TILNMFNLFQHSKLGVOVNLRIKLIILHETPPELYIGHGKMLSEFCQKQHEEGKK 411
Qy 303 SDKHPEY--W-----DTAIFLTRODLG--GATTCDTLGMADVGTMCDPKRSQVIED 350
Db 412 NDHLEMTNNGEDMTSVDAAILITRKDFCVHKDEPCDTVGIAVLSGCMSEKRCIIAED 471
Qy 351 DGLPSAFTTAHELGHVFNPHDN-VKVEEVEFGKLRANHMSPPLIQIDRAN--PWSACS 407
Db 472 NGNLNLAFTAHMGHNMGINHDNDHPSCAD-----GLHIMSGEWIKQNLGDVSWSPCS 525
Qy 408 AALITDFLDGSHGDCDLLD---QPSKPISELPELPGASYTLLSQCELAFGVSGKPCPYM 464
Db 526 KEDLERFLSKASNCLLQTNQSVNSVMVPSKLPQMTVTADEQCOILFGPLASFCQEQH 585
Qy 465 --CTLKWTGKAGQWQVOTRHF PWADGTCGEGKLCGLKAGCVERHNLNKRHVRVGSNAKW 522
Db 586 VICTGLWC--KVEGEKERTKLDPPMDGTDCLDGLKCKWCKAGECTSRTSAPER-LAGENSLW 642
Qy 523 DPGFCSRTCGGVQLARRQCTNPTPANGKYCEGVYRVKRSNLEPCSSASGSKSPREE 582
Db 643 ---SPCSRTCSAGISSREKRC--PGLDSEARDCNCPKQYRICENPPCPAGLPG--FDM 695
Qy 583 QCEAFNGYHNSTNRLTLAVAWPKYSVGSPPRDK-CKLICRANGTGYFYVLAPKVVVDGTL 641
Db 696 QCOAYSVRTSSPKHIL-----QMQAVLDEEPCALFCSVPKGEQPIILSEKVMGDTSC 748
Qy 642 SPDSTVCVQGGKCIKAGCDGNLGSKKRFDKCGVCGDNKSKKVTGLTKPM-HGYNFV 700

```
Db 749 GYGLDLCANGRCQKVGCDGLLGLSLAREDHGVCNGNGKCKIIKGDNFHTRGAGYVEVL 808
QY 701 AIPAGASSIDIRORYKGLIGDNDYALKNSOGKYLINGIPVVSAREEDLVKSLRLYS 760
Db 809 VIPAGARIKVVEEK-----FAHSYLAURDA-GKOSINSWKI-----EHS 848
QY 761 G-----TGTA-----ESLAQRSPLEPLTVEYL-----SVGKMTPPRVRYSFYLPKE 803
Db 849 GAFNLAGTTHVYVRGLWEKISAKGPTTAPLHLVLLVFDQDNYG-----LHYEYTPSD 902
QY 804 P-REDKSHKDP-----RGPVSLHNSVLSL-----829
Db 903 PUPENQSSKAPEPLFMWTHTSWEDCATCGGGERKTTSCTKINSKINISVDNECKKYL 962
QY 830 ---NOVEQPDPRP-PARVWAGSWGPCSGSLQKRAVDRCGSGAGORTVPACD-----AA 881
Db 963 KPEQIRKNEQPCOTRWMTWTWPCSTCKGQSGROVACTQQLSNGTLIRARERCIG 1022
QY 802 HRPVETQAC-GEPCPT-WELSAWSPCKSGCGRQFRRSLKCVGHGKLLARDQCNLHRKP 939
Db 1023 PKPASAQRCEGDCMVTNEAGWSEFSVKCGKGIHRTVRCNT-----PRKKCVLSTRP 1076
QY 940 QELDFC 945
Db 1077 REAEDC 1082

RESULT 11
Q8WXS8 PRELIMINARY; PRT; 1223 AA.
AC Q8WXS8;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE A disintegrin-like and metalloprotease with thrombospondin type 1
DE motif 14 precursor.
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638061; PubMed=11779638;
RA Bolz H., Ramirez A., von Brederlow B., Kubisch C.;
RT "Characterization of ADAMTS14, a novel member of the ADAMTS
RT metalloproteinase family";
RL Biochim. Biophys. Acta 1522:221-225(2001).
RL EMBL; AF38666; AAL40229.1; -
DR InterPro; IPR002870; Rep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Rep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Signal; Integrin; Protease; Metalloprotease.
FT SIGNAL 1 22 POTENTIAL.
SQ SEQUENCE 1223 AA; 133871 MW; 36394AC4D92F17DF CRC64;

Query Match 21.78; Score 1122; DB 4; Length 1223;
Best Local Similarity 29.98; Pred. No. 5e-85;
Matches 306; Conservative 122; Mismatches 379; Indels 216; Gaps 40;

QY 45 PEDSDQGLIFQITAFQEDFYHLTPDAQFLAPAFSTEH-----LGVPLQGLTGSSD 97
Db 98 PGRVGRHSLYENVTVFGKELHLRLRPNRLVVPVGSVSEWQEDPTELFRLR-----149
QY 98 LRRCFYSDVNAEPDPSFAVNSLGGGLGAFGYRGAETVISPPLNAPAPAAQRNSQ-AHL 156
Db 150 -QECVYTGVTGMPGAAVAISNCNCDGLAGLIRTDSTDFIEPLERQ-----QKEASGRTHV 205
```

```
QY 157 LQRRGVP-----GPGSDPTSRGCVASGNPAILLRALDPYKPRAGFGESESRRRSSKAKR 212
Db 206 YVREAVQOEWAEPDGDGLHNE-AFGLDLPNLLGLVGDO-----LGTERRKRHAKPGS 258
QY 213 FVSIPRVETLVVADESMVKFHGAD-LEHYLLTLLATAARLYRHPHSILNPINTVVVKVLL 271
Db 259 Y-----SIEVLLVVDSDVVRFHGKEHVQNVLTLMIVDEIYHDESGLGVHINIALVRLIM 313
QY 272 LRDRDSGPV-TGNAALTLENFCAMQKLNKYSKHIPEYWDTAALTFTRODLGATTCDFL 330
Db 314 VGYRSLSLIERGNPSSLEQVCRWAHSOQRPDPSHAHHHDHVVFLTRQDF-----GPS 367
QY 331 GMADVGTMCDPKSCSVIEDDGLPSAFTAHGLGHVFNPHD---NVKYCEEVFGKIRAN 387
Db 368 GYAPVTGCMCHPLKSCALNHEDGFSSAFVTAHETGHVMEHDGQGGCAGDETSLSG-----423
QY 388 HMMSTLIQIDRANPNWASCSAAIITDPLDSGHGDCILDDQSPKI-SLPEDLPGASVTLG 446
Db 424 -VMAPLVQAAAFHFRHWSRCSKLELSRYLPS--YDCLDDPDPAPWPQPELPGINYSMDE 480
QY 447 QCELATGVGSKPC---PYMOYCTKLWCTGAKGOMYQCTRHFPWADGTSCGEGKLCCKGA 503
Db 481 QCRFDGSGYQTCIAFRTEPECKQLWCS-HPDNPYFCKTKGPPLDGTECAPGKWCFKGH 539
QY 504 CVERHNLNKHVRVDGSAKWDPYGPCSRTCGGVQLARROCTNPTTPANGKYCEGVRYKYR 563
Db 540 CIWKSPEQIYQDGGWSSWTKFGCSRSRSGGVRSRSCNNPSPAYGRPCUGLPMFEYQ 599
QY 564 SCNLEPCPSASGKSFREOCEAFNGY-----NHSNRLTLAVAVVPKYSGVSPDKCK 617
Db 600 VCNSEECPTGY--EDFRAQCAKRNYSYVHQNAKHS-----WVP-YEPDQDAQKCE 647
QY 618 LICRANGTYFYVLAQVVDGTLCS-PDSTSVQVQCKIACGDNGLGSKKRFKDCGCVG 676
Db 648 LICQSDATGVDWFMNQVVDGTRCSYRDPYSVCARGECPVGVGDKVGSNKAADKCGVCG 707
QY 677 GDNKCKKYVTGLFTK--PMHGYNFVVAIPAGASSIDIRQGVKGLIGDNDYALKNS-QG 733
Db 708 GDNSHCRTVKGTLGKASKQAGALKVQIPAGARHIQI-----EALSKSPHRIVVKVQVVG 762
QY 734 KYLLN--GHFVVSAREDLVVKGSLRLYSQGTCTAVESLQASRPILPLTVEVLSVGKMT 791
Db 763 SFILNPKGK--EATSRTEFAMG-LEWEDAVEDAKESLTSGLTSPLEPAIALAL-----811
QY 792 PRVRYSFYLPKEPREDKSSHKDPGRPS-----VLHNSVLSL--SNQV--EQPDRPP 840
Db 812 -----PTEGGPRSLAYKYVIHEDLLPLIGSNVLLLEEMD---T 848
QY 841 ARWVAGSWGPCSGSLQKRAVDRC-----867
Db 849 YEWALKSWAPCSKACCGGIQFTYKGCRRRRDHMHVORHLCDDHKRKPPIRRRCNQHPCSQ 908
QY 868 -----GSAGQRT-----VPACDAARHPVETQACG-----EPC 894
Db 909 PVWVTEWAGCSRSCGKLGVTQIRGICQLPLSLNGTHKVPKACACADGDRPEARPCLRVPC 968
QY 895 PT-WELSAWSPCKSGCGRQFRRSLK-----VHGGRLLARDQCNLHRKPOELDFCVL 947
Db 969 PAOWRLGAWSQCSATCGEIQOQVQVVCRTNANSLGH-----CEGDR-PDTVQVCSL 1018
QY 948 RPC 950
Db 1019 PAC 1021

RESULT 12
Q8TEY8 PRELIMINARY; PRT; 1159 AA.
ID Q8TEY8;
AC Q8TEY8;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
```

```
DE ADAMTS14.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21839041; PubMed=11741898;
RA Collige A., Vandenbergh I., Thiry M., Lambert C.A., Van Beeumen J.,
RA Li S.W., Prockop D.J., Lapiere C.M., Nussgens B.V.;
RT "Cloning and characterization of ADAMTS-14, a Novel ADAMTS Displaying
RL High Homology with ADAMTS-2 and ADAMTS-3.";
DR J. Biol. Chem. 277:5756-5766(2002).
DR EMBL; AF366351; RAL79814.1;
SQ SEQUENCE 1159 AA; 127336 MW; A5B130149BF7F34 CRC64;

Query Match 21.7%; Score 1120.5; DB 4; Length 1159;
Best Local Similarity 30.0%; Pred. No. 6.2e-85;
Matches 307; Conservative 123; Mismatches 380; Indels 213; Gaps 41;

QY 45 PEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAFSTEH-----LGVPLQGLTGSSD 97
DB 31 PGRVGRHSLYFNVTGKELHLRLRNRLVPGSSVQWQEDFRELFRQLR----- 82
QY 98 LRCFTSGDVNAEDPSFAAVLGGGLRGAGFYRGABYVISPPLNPASAPAAQNSQG-AHL 156
DB 83 -QECVVTGGVTGMPGAATAISNCDGLAGLIRTDSTDFIEPLERGO--QKEASGRTHV 138
QY 157 LQRRGVP---GPGSDPTSRCGVASGNWPAILRALDPYKPRAGFESRRSRGRKR 212
DB 139 YVREAVQOEWAEPDGLHNE-AFGLDLPNLLGLVGDO---LGDTERKRHRKPGS 191
QY 213 FVSIPRYVETLVVADESVMKFGHAD-LEHYLLTLTAARLYHPSILPINIVVVKLL 271
DB 192 Y-----SIEVLLVDDSVVRFGHEVQVYVLTLMNIVDEIYDESLGVHINIALVRLIM 246
QY 272 LRDRDSGPKV-TGNAALTILRNFCAMQKLNKYSKDHPEYWDITILFTRODLGATTCDTL 330
DB 247 VGYRQSLSLIERGNPSRSLQVCRWAHSQORQDPSSHAHHDHVFLTRQDF-GPSGMQ-- 303
QY 331 GMADVTGTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHD---NVKVEEVEFKLRAN 387
DB 304 GYAPVTGCMUPLRSCALNHEDGESSAFVIAHETGHVGMHEHVGNGCADETSLGS---- 359
QY 388 HMSPTLIQIDRANPWSACSAIITDLDLSDGDCLLDQPSKPI-SLPEDLPFCASYTLQ 446
DB 360 -VMAPLVQAAAFHFRHSRCSKLSRLPS--YDCLDDPDPAPWQPPPELPGINYSMDE 416
QY 447 QCELATFGVSGKPC---PYMQYCTKLINCTGKAKQGMVQOTRHPFPWADGTSCGEGKLC 503
DB 417 QCRFDGSGYQTCIAFTFTPECPKQWCS-HPDNPYFCKTKGPPLDGTETCAPGKWCFKGH 475
QY 504 CVERHNLNKHVRDGNWAKMDPYGPCSRTCGGGVQLARRQCTNPTPANGKCYCEGVKRYR 563
DB 476 CIWKSPEQTFQDGGHSSWTGFCSCSRSCGGVYRSRSCNPNPSPAYGRCPLGPMFEYQ 535
QY 564 SCNLEPCPSSAGSKSFREEQCEAFNGY-----NHSTNRLTLAVAWPKYGVSPRDKCK 617
DB 536 VCNSECEPGTY--EDFRAQCAKRNYSYVHQNAKHS-----WVP-YEPDDDAQKCE 583
QY 618 LICRANGTGYFYVLAPKPVVDGTLCS-PDSTVCVQCKCIKAGCDGNLGSKKRFDKCGVCG 676
DB 584 LICOSADTGVDFPMNOVHDGTCTCSRDPSYVCARGCEVPVGCDEKVEGSMKADKCGVCG 643
QY 677 GDNKSKKVTGLTK--PHGYFNVAIPAGASSIDIRQYKGLTGDNDYLAALKNS-OG 733
DB 644 GDNSHCETVGTGLGKASKQAGALKVQIPAGARHIQI-----EALKSPHRIVKNQVGTG 698
QY 734 KYLLN--GHEVSAVERDLVYKSGLLRYSGTGTAVESLQASRPILEPLTVELSVGKMTP 791
DB 699 SFILNPKGK---EATSRFTTAMG-LEWEDAVEDEKESLTKSGPLPEAIALA----- 747
```

```
QY 792 PRVRYSTYLPKEPREDKSSHPKDPGRPS-----VLHNSVLSL--SNQV--EOPDDRRPP 840
DB 748 -----PTEGGPRSSLAYKYVIHEDLLPLIGSNVLLLEMD---T 784
QY 841 ARWVAGSWGPCASCSGSLQKRAVDCH----- 867
DB 785 YEWALKSWAPCSKACGGGIGFTYGCRRRRDHHVQRHLCHDKKPKPIRRRCNQHPCSQ 844
QY 868 -----GSAGQRT-----VPACDAARHPVETQACG-----EPC 894
DB 845 PVWVTEWAGACSRSCGKLGVTQTRGICQLLPLSNGTHKVPKACAGDRPEARREPCLRVPC 904
QY 895 PT-WELSAWSPCKSCGGRGQRRLSKC-----VHGGRLLARDQCNLHKKPOELDFCVL 947
DB 905 PAQWRLGAWSCSACGEGIQORVQVCRTNANSIGH-----CEGDR-PDTVQVCSL 954
QY 948 RPC 950
DB 955 PAC 957

RESULT 13
Q8TE55 PRELIMINARY; PRT; 1223 AA.
AC Q8TE55;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metalloprotease-disintegrin protease.
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002);943.1;
DR EMBL; AJ345098; CAC87943.1;
KW Integrin; Protease.
SQ SEQUENCE 1223 AA; 133925 MW; D585B6593977ED15 CRC64;

Query Match 21.7%; Score 1118; DB 4; Length 1223;
Best Local Similarity 29.8%; Pred. No. 1.1e-84;
Matches 305; Conservative 123; Mismatches 379; Indels 216; Gaps 40;

QY 45 PEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAFSTEH-----LGVPLQGLTGSSD 97
DB 98 PGRVGRHSLYFNVTGKELHLRLRNRLVPGSSVQWQEDFRELFRQLR----- 149
QY 98 LRCFTSGDVNAEDPSFAAVLGGGLRGAGFYRGABYVISPPLNPASAPAAQNSQG-AHL 156
DB 150 -QECVVTGGVTGMPGAATAISNCDGLAGLIRTDSTDFIEPLERGO--QKEASGRTHV 205
QY 157 LQRRGVP---GPGSDPTSRCGVASGNWPAILRALDPYKPRAGFESRRSRGRKR 212
DB 206 YVREAVQOEWAEPDGLHNE-AFGLDLPNLLGLVGDO-----LGDTERKRHRKPGS 258
QY 213 FVSIPRYVETLVVADESVMKFGHAD-LEHYLLTLTAARLYHPSILPINIVVVKLL 271
DB 259 Y-----SIEVLLVDDSVVRFGHEVQVYVLTLMNIVDEIYHDSGLGVHINIALVRLIM 313
QY 272 LRDRDSGPKV-TGNAALTILRNFCAMQKLNKYSKDHPEYWDITILFTRODLGATTCDTL 330
DB 314 VGYRQSLSLIERGNPSRSLQVCRWAHSQORQDPSSHAHHDHVFLTRQDF-----GPS 367
QY 331 GMADVTGTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHD---NVKVEEVEFKLRAN.387
```

```
Db 368 GYAPVTGCHPLRSCALNHEDGFSAFVIAHETGVLGMEHDGCGCADETSLGS----- 423
Qy 388 HMSPTLIQIDRANPWSAIAITDFLDLSGHGCDLDDQSKPI-SLPEDLPGASYTLQ 446
Db 424 -VMAPLVQAARFHWSCSLKSLRYLPS--YCLDDDFDPAWPPFELPLGINSMD 480
Qy 447 QCELAFGVSGKPC---PMQYCTKLWCTGKAKGQMVQCTRHFPWADGTSCGEGKLCIKGA 503
Db 481 QCRDFGSGYOTCLAFRTFEPCKLWCS-HPDNPFYCKTKKGLPPLDTECAPKWCFKGH 539
Qy 504 CVERHNKLRVDSWAKWDYGCSTTCGGVGLARQCTNTPPANGKYCEGVRYK 563
Db 540 CIWKSPEQTYQDGGWSWTKFGSCSRSCGGVRSRSCNNPSPAYGGRPCIGPMEYQ 599
Qy 564 SCNLEPCPSSAGSKSFRECEAFNGY-----NHSTNRLTLAVAVWPYKSGVSPRDKCK 617
Db 600 VCNSECPGTY--EDFRAQCAKRNYSYVHQAHS-----WYP-YEPDDDAQKCE 647
Qy 618 LICRANGTYGYVLAPKVVDGTGCS-PDSTSVQVQKICAKGCDGNLGSKKRFDKCGVCG 676
Db 648 LICQADTGDVFMQNVHDGTRCSYRDYSVCAARGCEVPVGCDEKVGSKHADDKCGVCG 707
Qy 677 GDNCKKVTGLFTK--PMHGYNFVVAIPAGASSIDTRQGYKGLIGDNYLAKNS-QG 733
Db 708 GDNHCRVTGTLGKSKQAGALKLVQIPAGARHIQI-----EALSKSPHRIVVKQVTG 762
Qy 734 KYLLN--GHFVVSVERDLVVKGLLYSGTGTAVESLOASRPTLEPLTVEVLSVGKMT 791
Db 763 SFILNPKG---EATSTFTAMG--LEWEDAVEDAKESLKTSGPLPEAIAL----- 811
Qy 792 PRVYSYLPKEPREDKSSHPKDRGPS-----VLHNSVLSL--SNQV--EQPDRPP 840
Db 812 -----PTEGGPSLAVKYVIHEDLLPLIGSNVLLLEMD---T 848
Qy 841 ARWAGSWGPCSCSGSLQKRAVDCR----- 867
Db 849 YEWALKSWAPCSKACGGGIRFTYGCRRRRDHMHVHRHLDCHKRPKIRRRSNQHPCSQ 908
Qy 868 -----GSAGORT-----VPACDAARHPVETQAC-----EPC 894
Db 909 PVWTEBWGACRSYKGLGVQTRIGIQCLLPLSNCTHKVMPAKACAGRPEARRPCLRVPC 968
Qy 895 PT-WELSAWSPCKSCSGRQFORSKLC-----VHGGRLLARDQCNLHKPKQELDFCVL 947
Db 969 PAOWRLGAWSCSATCGEGIQQRVQVVCRTNANSLGH-----CEGDR-PDTVQVCSL 1018
Qy 948 RPC 950
Db 1019 PAC 1021

RESULT 14
Q96L37
ID Q96L37 PRELIMINARY; PRT: 1427 AA.
AC Q96L37;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX PubMed=11557746;
RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E.,
RA Fujikawa K.;
RT "Structure of von Willebrand factor-cleaving Protease (ADAMTS13), a
RT Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura";
```

```
J. Biol. Chem. 276:41059-41063(2001).
RL EMBL: AY055376; AAL17652.1;
DR MEROPS; M12.241;
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR000130; Zn_MTPeptidse.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00090; tsp.1; 4.
DR PROSITE: PS0215; ADAM_MEPPO; 1.
DR PROSITE: PS00092; TSPI; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Protease; Signal.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; EB1BC3AABCI4442 CRC64;

Query Match 17.3%; Score 893.5; DB 4; Length 1427;
Best Local Similarity 29.2%; Pred. No. 1e-65;
Matches 250; Conservative 108; Mismatches 347; Indels 151; Gaps 34;

Qy 174 CGVASQ-WNP-----AILRALDPY-----KPRRAGGESRRSRSGAKRF 213
Db 21 CGFLGCGWPSHFQSCIQALEPQAVSSYLSPGAPLKPSPSPGFQQRQOR-----RA 75
Qy 214 VSTPRYVETLVVADESMVKEFHGADLEHLLTLATAARLYRHPSTILNPINIVVVKVLLR 273
Db 76 AGGILHLELLVAVDPVFOAHOEDTERVLTNLTGAEALLRDPDSLGAOFRVHLVRMVL 135
Qy 274 DRDSGKVTGNAALTLRNFCWOKLKYSDKHPEYWDTAIFLTRQDL-CGATTCDTLGM 332
Db 136 EPEGAPNITANLTSSLLSVCGSQINPEDDTPGHADLVLYITRFDLELDPGNRQVRG 195
Qy 333 ADVGTMCDPKRSCSVIEDDGLPSATTAHELGHVNMHPDNVVKVEEFGKRAHNMSP 392
Db 196 TOLGGACSPWTSCLITDGTDLGVITIAHEIGSFLEHGDGAPG-----CGFSGHVMA 251
Qy 393 TLIQIDRANP-----WSACSAIITDFLDSGHGDCLLD-----OPSKPISLPEDLP 443
Db 252 -----DGAAPRAGLAWSPCSRRLQLLSLSAGRARCVWDPPRPPQGSAGHPDPAQCL 306
Qy 444 LSQCELAFVGSKPCPY-----MOYCTKLWCTGKAKGQMVQCTRHFPWADGTSCGEGK 499
Db 307 ANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLVPLLDGTGECVEN 366
Qy 500 LKAGACVERHNLNK-HRVDSWAKWDYGCSTTCGGVGLARQCTNTPANGKYCEGV 558
Db 367 SKGRCSRLVELPIAAVHGRWSSWGPSPCRSCGGVVTTRRQCNRPAPFGGRACVA 426
Qy 559 RVKRSCLNLEPCPSSASGSKSFRECEAFNG-----YNHSTNRLTLAVAW---V 604
Db 427 DLQAEACNTQACEKQ--LEFMSQOCARTDQGPLRSSPGASFYH-----WGA 474
Qy 605 PRYSGVSPDKCLICRANGTYGYVLAPKVVDGTGCLSP-----DST-SVCVCGKCIAG 658
Db 475 PHSQGD---LCRHMCRAIGESFIMKRGDSFLDGTFRMPSGPRREDGTSLCLVSGSCRT 531
Qy 659 CDGNLGSKKRPDKCGVCGDNKSKKVTGLFT--KPMHGYNFVVAIPAGASSIDTRQRY 716
Db 532 CDGRMDSQOVWDRQVCGDNDSTCSPRKGSFTAGRAREYVTLVTP--NLTSVYI--ANH 588
Qy 717 KGLIGDDNYLALKNSQGYLLNGHFVVS--AVERDLVVKGSL-LRYSGTGTAVESLOAS 773
Db 589 RPLF---THLAVRIG-GRYVAVAGKMSISPNTTYPSSLLEDGRVEYRVALTEDRLPLEIR 644
Qy 774 ---PILEPLTVEVL-----SVGKMTPPRVYSFYLPKPEPREDKSSHDPKDPSPVLHNS 825
Db 645 IWGPLQEDADIQVYRRYGEEYGNLTRPDITFTYQPK----- 681
Qy 826 LSLSNQVQEPDRPPARWV-AGSWGPCSCSGSLQKRAVDCRGSAGORTVPA--CDAAH 882
Db 682 -----PQAWVVAARVGPSCVSCAGLRWVNYSCLDQARKELVETVQVQSG 728
Qy 883 R-PVETQACG-EPDPT-WELSAWSPCKSCSGRQFORSKLCVGHGGRLL-----AR 934
```


Sat May 17 11:00:00 2003

Db 729 OPPANPEACVLEPCPPYNAVGFPCSGGLRERVRVCEAOGSLKTLPPARCRAG 788
QY 935 LHRKPQELDFCVLRPC 950
Db 789 AOPVALETCPNPQC 804

RESULT 15

O9GL54 PRELIMINARY; PRT; 269 AA.
AC O9GL54;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Aggrecanase-2 (Fragment).
OS ADAMS-11.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Goad D.L.; Goad M.E.;
RT "Molecular cloning of lapine aggrecanase-2 (ADAMS-11) from articular chondrocytes."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBAJ databases.
DR EMBL; AF317415; AAG33062.1; .
DR HSSP; O9PW35; 1BUD.
DR MEROPS; M12.225; .
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 1.
DR SMART; SM00209; TSPI; 1.
DR PROSITE; PS0215; ADAM_WEPPO; 1.
DR PROSITE; PS0092; TSPI; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;

Query Match 16.1%; Score 830; DB 6; Length 269;
Best Local Similarity 54.7%; Pred. No. 1.8e-61;
Matches 146; Conservative 37; Mismatches 74; Indels 10; Gaps 3;

QY 329 TLGMADVGTMCDPKRSCTVEDDGLPSAFTTAHELGHVENMPHDNVKCEEVFGKLRANH 388
Db 1 TLGMADVGTTCSPERSCAVIEDDGLHAFTVAHEIGHLLGLSHDSDKFCCEENFGSTEDKR 60
QY 389 MMSPTLTIOIDRANPWSACSAIITDFLDGSHGDCLLDQPSKIPSLPDLPGASYTLQQC 448
Db 61 LMSILTSIDASKPWSKCTSAITTEFLDDGHGNCLLDVPKQILGPEELPGOTYDATQOC 120
QY 449 ELAFGVGSKPCPYMYQCTKLWCTGKAGQMVQCTRHFPWADGTSCEGKCLKGACVER- 507
Db 121 NLTFGPEYTVCPGMDVCARLWCVAVRQGMVCLTKLPALVEGTPCGKGRICLQKCVDKT 180
QY 508 ----HNLNKHRVDGNAKWDPTGCSRTCTGGVQLARROCTNPTPANGKGYCEGVRYK 563
Db 181 KKKYISTSSH---GNWGSWGPWGQCSRCGGVQFAIRHCNNPAPRNSGRYCTGKRIYR 237
QY 564 SCNLEPCPSSAGSKSPFEQCEAFNGY 590
Db 238 SCSVTPCP--ANGKSPHEQCEAKNGY 262

Search completed: May 9, 2003, 15:23:35
Job time : 54 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:23:41 ; Search time 56 Seconds
(without alignments)
1561.150 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLTILTAFAAGTAGGSE.....DQCNLHRKPOELDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5162	100.0	950	10	US-09-965-631-4
2	4248.5	82.3	823	9	US-10-163-316-2
3	2486.5	48.2	967	12	US-10-105-929-2
4	2485.5	48.1	950	10	US-09-321-987B-4
5	2485.5	48.1	968	9	US-10-163-316-7
6	2274	44.1	727	9	US-10-097-597-1
7	2274	44.1	727	9	US-10-097-580-1
8	2274	44.1	727	10	US-09-445-023A-1
9	2273	44.0	727	9	US-10-097-597-12
10	2273	44.0	727	9	US-10-097-580-12
11	2273	44.0	727	10	US-09-445-023A-12
12	2115	41.0	905	10	US-09-918-171A-9
13	1965	38.1	1629	10	US-09-972-467-2
14	1919	37.2	367	10	US-09-965-631-6
15	1914	37.1	837	9	US-10-174-590-352
16	1914	37.1	837	9	US-10-176-758-352
17	1914	37.1	837	9	US-10-175-737-352
18	1914	37.1	837	9	US-10-173-706-352
19	1914	37.1	837	9	US-10-175-738-352

```
20 1914 37.1 837 9 US-10-175-752-352 Sequence 352, App
21 1914 37.1 837 9 US-10-176-482-352 Sequence 352, App
22 1914 37.1 837 9 US-10-176-757-352 Sequence 352, App
23 1914 37.1 837 9 US-10-176-913-352 Sequence 352, App
24 1914 37.1 837 9 US-10-180-552-352 Sequence 352, App
25 1914 37.1 837 9 US-10-180-557-352 Sequence 352, App
26 1914 37.1 837 9 US-10-173-700-352 Sequence 352, App
27 1914 37.1 837 9 US-10-174-572-352 Sequence 352, App
28 1914 37.1 837 9 US-10-174-579-352 Sequence 352, App
29 1914 37.1 837 9 US-10-174-582-352 Sequence 352, App
30 1914 37.1 837 9 US-10-174-588-352 Sequence 352, App
31 1914 37.1 837 9 US-10-175-739-352 Sequence 352, App
32 1914 37.1 837 9 US-10-175-740-352 Sequence 352, App
33 1914 37.1 837 9 US-10-175-743-352 Sequence 352, App
34 1914 37.1 837 9 US-10-176-488-352 Sequence 352, App
35 1914 37.1 837 9 US-10-176-492-352 Sequence 352, App
36 1914 37.1 837 9 US-10-176-747-352 Sequence 352, App
37 1914 37.1 837 9 US-10-176-750-352 Sequence 352, App
38 1914 37.1 837 9 US-10-176-985-352 Sequence 352, App
39 1914 37.1 837 9 US-10-176-987-352 Sequence 352, App
40 1914 37.1 837 9 US-10-176-991-352 Sequence 352, App
41 1914 37.1 837 9 US-10-176-992-352 Sequence 352, App
42 1914 37.1 837 9 US-10-176-993-352 Sequence 352, App
43 1914 37.1 837 9 US-10-184-658-352 Sequence 352, App
44 1914 37.1 837 9 US-10-173-695-352 Sequence 352, App
45 1914 37.1 837 9 US-10-173-697-352 Sequence 352, App
```

ALIGNMENTS

```
* RESULT 1
US-09-965-631-4
; Sequence 4, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020115842A1el Human Proteases and Polypeptides Enco
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-965-631-4
```

```
Query Match 100.0%; Score 5162; DB 10; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MLLGLTILTAFAAGTAGGSEPEREVVVRILDPDINGRYYWRGSDSDQGLIFQITAF 60
Db 1 MLLGLTILTAFAAGTAGGSEPEREVVVRILDPDINGRYYWRGSDSDQGLIFQITAF 60
QY 61 QEDFYHLTPDQAQFLAPAFSTEHGLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
Db 61 QEDFYHLTPDQAQFLAPAFSTEHGLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
QY 121 GGLRGAFYRGAEYVTSPLPNASAPAAQRNSOGAHLQRGYPGGPSGDPTRSCVAGSW 180
Db 121 GGLRGAFYRGAEYVTSPLPNASAPAAQRNSOGAHLQRGYPGGPSGDPTRSCVAGSW 180
QY 181 NPAIRALDPYKPRAGGESRRRSRAKRFVSIPIRYVETLVVADESMVKFHGADLEH 240
Db 181 NPAIRALDPYKPRAGGESRRRSRAKRFVSIPIRYVETLVVADESMVKFHGADLEH 240
QY 241 YLLTLLATAARYRHPISILNPINIVVVKLLLRDSDSGPKVTGNAALTILRNFCAWQKILN 300
```

```
Db 241 YLLTLLTAARLYRHPISILNPIVNVKVLRLDRSGPKYGTNAALTLENFCAWKKLN 300
QY 301 KVSXKHPEYWDTAIFTRQDLGATTCDTLGMADVGTWCDPKRSCSVIECDGLPSAFTTA 360
Db 301 KVSXKHPEYWDTAIFTRQDLGATTCDTLGMADVGTWCDPKRSCSVIECDGLPSAFTTA 360
QY 361 HELGHVFNPHDNVKEVFGKLRANHMMSPTLTIQIDRANPWSACSAAITDFLDSGHG 420
Db 361 HELGHVFNPHDNVKEVFGKLRANHMMSPTLTIQIDRANPWSACSAAITDFLDSGHG 420
QY 421 DCLLDQSPKPSILPEDLPASVYTLSSQCELAFGVSKPCPYMYQCTKLWC?CKAKGQWVC 480
Db 421 DCLLDQSPKPSILPEDLPASVYTLSSQCELAFGVSKPCPYMYQCTKLWC?CKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYVDGWSAKWDYPCGSETCGGVQVLAR 540
Db 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYVDGWSAKWDYPCGSETCGGVQVLAR 540
QY 541 ROCTNPTPANGGKYGCVGRVYKRSNLEPCPSSASGKSPREOCBAFNGYNHSTNRLTLA 600
Db 541 ROCTNPTPANGGKYGCVGRVYKRSNLEPCPSSASGKSPREOCBAFNGYNHSTNRLTLA 600
QY 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGFTLCSPDSTSVCVQCKCIKAGCD 660
Db 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGFTLCSPDSTSVCVQCKCIKAGCD 660
QY 661 GNLSKKRFDKCGVCGDGNKSKKVTGLTKPMHGFNFVVAIPAGASSIDIRQGYKGLI 720
Db 661 GNLSKKRFDKCGVCGDGNKSKKVTGLTKPMHGFNFVVAIPAGASSIDIRQGYKGLI 720
QY 721 GDDNYLALNKSQGYLLNGHFVYSAVERDLVWKGSLRYSGTGTAVESLQASRPILPLT 780
Db 721 GDDNYLALNKSQGYLLNGHFVYSAVERDLVWKGSLRYSGTGTAVESLQASRPILPLT 780
QY 781 VEVLVSGKMTPPRRVYSFYLPKPEPRDKSSHPRGKDPGRPSVLHNSVLNSNOVEQDDRRP 840
Db 781 VEVLVSGKMTPPRRVYSFYLPKPEPRDKSSHPRGKDPGRPSVLHNSVLNSNOVEQDDRRP 840
QY 841 ARWVAGSWGPCASCSGGLQKRAVDCRGSGAGORTVPACDAHRPVEVTOACGPECPITWELS 900
Db 841 ARWVAGSWGPCASCSGGLQKRAVDCRGSGAGORTVPACDAHRPVEVTOACGPECPITWELS 900
QY 901 AWSPCSKSGRGRFQRRSLKCVGHGRLARQDQCNLHRKPJELDFCVLRPC 950
Db 901 AWSPCSKSGRGRFQRRSLKCVGHGRLARQDQCNLHRKPJELDFCVLRPC 950
```

```
RESULT 2
US-10-163-316-2
; Sequence 2, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI01-075PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-316-2
```

```
Query Match 82.3%; Score 4248.5; DB 9; Length 823;
Best Local Similarity 97.1%; Pred. No. 6.9e-295;
Matches 792; Conservative 0; Mismatches 1; Indels 23; Gaps 1;
```

```
QY 1 MLLGLILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
Db 1 MLLGLILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
QY 61 QEDFYHLTTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
Db 61 QEDFYHLTTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
QY 121 GGLRGAFYGAEVVISPPLNAPAAORNSOGAHLQRRGVPGGSPGDPSTSCGVASGW 180
Db 121 GGLRGAFYGAEVVISPPLNAPAAORNSOGAHLQRRGVPGGSPGDPSTSCGVASGW 180
QY 181 NPALLRALDYPKPRRAGFESRRSRGRKRFVSPRYVETLWVADESMVKFHGADLEH 240
Db 181 NPALLRALDYPKPRRAGFESRRSRGRKRFVSPRYVETLWVADESMVKFHGADLEH 240
QY 241 YLLTLLTAARLYRHPISILNPIVNVKVLRLDRSGPKYGTNAALTLENFCAWKKLN 300
Db 241 YLLTLLTAARLYRHPISILNPIVNVKVLRLDRSGPKYGTNAALTLENFCAWKKLN 300
QY 301 KVSXKHPEYWDTAIFTRQDLGATTCDTLGMADVGTWCDPKRSCSVIEDDGLPSAFTTA 360
Db 301 KVSXKHPEYWDTAIFTRQDLGATTCDTLGMADVGTWCDPKRSCSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNPHDNVKEVFGKLRANHMMSPTLTIQIDRANPWSACSAAITDFLDSGHG 420
Db 361 HELGHVFNPHDNVKEVFGKLRANHMMSPTLTIQIDRANPWSACSAAITDFLDSGHG 420
QY 421 DCLLDQSPKPSILPEDLPASVYTLSSQCELAFGVSKPCPYMYQCTKLWC?CKAKGQWVC 480
Db 421 DCLLDQSPKPSILPEDLPASVYTLSSQCELAFGVSKPCPYMYQCTKLWC?CKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYVDGWSAKWDYPCGSETCGGVQVLAR 540
Db 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYVDGWSAKWDYPCGSETCGGVQVLAR 540
QY 541 ROCTNPTPANGGKYGCVGRVYKRSNLEPCPSSASGKSPREOCBAFNGYNHSTNRLTLA 600
Db 541 ROCTNPTPANGGKYGCVGRVYKRSNLEPCPSSASGKSPREOCBAFNGYNHSTNRLTLA 600
QY 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGFTLCSPDSTSVCVQCKCIKAGCD 660
Db 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGFTLCSPDSTSVCVQCKCIKAGCD 660
QY 661 GNLSKKRFDKCGVCGDGNKSKKVTGLTKPMHGFNFVVAIPAGASSIDIRQGYKGLI 720
Db 661 GNLSKKRFDKCGVCGDGNKSKKVTGLTKPMHGFNFVVAIPAGASSIDIRQGYKGLI 720
QY 721 GDDNYLALNKSQGYLLNGHFVYSAVERDLVWKGSLRYSGTGTAVESLQASRPILPLT 780
Db 721 GDDNYLALNKSQGYLLNGHFVYSAVERDLVWKGSLRYSGTGTAVESLQASRPILPLT 780
QY 781 VEVLVSGKMTPPRRVYSFYLPKPEPRDKSSHPRGKDPGRPSVLHNSVLNSNOVEQDDRRP 840
Db 781 VEVLVSGKMTPPRRVYSFYLPKPEPRDKSSHPRGKDPGRPSVLHNSVLNSNOVEQDDRRP 840
QY 841 ARWVAGSWGPCASCSGGLQKRAVDCRGSGAGORTVPACDAHRPVEVTOACGPECPITWELS 900
Db 841 ARWVAGSWGPCASCSGGLQKRAVDCRGSGAGORTVPACDAHRPVEVTOACGPECPITWELS 900
QY 901 AWSPCSKSGRGRFQRRSLKCVGHGRLARQDQCNLHRKPJELDFCVLRPC 950
Db 901 AWSPCSKSGRGRFQRRSLKCVGHGRLARQDQCNLHRKPJELDFCVLRPC 950
```

```
RESULT 3
US-10-105-929-2
; Sequence 2, Application US/10105929
; Patent No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
```

;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 967
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-105-929-2

Query Match 48.2%; Score 2486.5; DB 12; Length 967;
Best Local Similarity 49.0%; Pred. No. 4.5e-169;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;
QY 1 MLLGILTLAFAGTAGSEPREVVPIRLDPDINGRRYVRGPEDSGQGLIFQITAF 60
Db LLLLLAALLAVSDALGRSEDEELVLP-ELE-----RAP---GHGTTTLRLHAF 81
QY 61 QDELYLHTLPDQAFLAPAFSTHGLG-----VPLQGLGSGSDLRRCFYSGDVNAEPDSF 114
Db DQQLDLELRPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFYGAAYVISPPLNAS---APAAQNSOGA-----HLLQ---RRGVPGG 165
Db AALSUCGVRGAFYLLGGEAYFTQPLPAASERLATAAGKEKPPAPLQFHLRLRNQGGDVG 196
QY 166 PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRRAGFSGSRSSRS 207
Db TCGVVDDEPRPTGKAETDEDEGTGEDEGPQWS-----PQDPALQGVGP-TGTGS 247
QY 208 GRAKRFVSPRYVETLVVADESVMKFGHADLEHYLLTLATAARLYRHPS:LNFINIVVV 267
Db IRKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFVAARLYKHPS-RNSVSLVVV 307
QY 268 KYLLLRDRSGKYVGNALTLRNCANQKLNKYSDKHPEYWDTAIFYHQDLGGATTC 327
Db KILVTHDEQKGPVTSNAALTLRNCANQKLNKYSDKHPEYWDTAIFYHQDLGGATTC 367
QY 328 DTILGMADVGTMDCKPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVVCVEEFGKLRAN 387
Db DTILGMADVGTMDCKPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVVCVEEFGKLRAN 427
QY 388 HNSPFLTIQIDRANPWSACSAIITDLDGSHGDCLLDQPKSPISLPEDLPASVTLTQQ 447
Db HNSPFLTIQIDRANPWSACSAIITDLDGSHGDCLLDQPKSPISLPEDLPASVTLTQQ 487
QY 448 CELAFGVGSKPCP-YWQYCTKLWCTGKAKQWVCOTRHPFADGTSCGEGKLCGLKACVE 506
Db COFTFGDSKHCPCDAASTCTLWCTGSGGLVLCQTKHFPWADGTSCGEGKLCGLKACVE 547
QY 507 RHNLNKH---RVDGSKWAKWDPYGSCRTCGGGVQLARRQCTNPTPANGSKYCEGVYKVR 563
Db KTD-RKHFTPHGSGWGMGPMGDCSRCTCGGGVQVYTMRECONVPKNGKCYCEGKRYR 606
QY 564 SCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRDKCLICRAN 623
Db SCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRDKCLICRAN 665
QY 624 GTCGYFVLAPKVDGTGLSPDSTSCVQKCKIACGDNGLSKKFKPCVCGGDNKCK 683
Db GTCGYFVLAPKVDGTGLSPDSTSCVQKCKIACGDNGLSKKFKPCVCGGDNKCK 725
QY 684 KYTGFTKPMHGYNFWAIPAGASSIDIRORYKGLIGDNYLALKNSQKYLINGHFTVV 743
Db KYTGFTKPMHGYNFWAIPAGASSIDIRORYKGLIGDNYLALKNSQKYLINGHFTVV 785
QY 744 SAVERDLVVKGLSYGTGTAVESLOASRPILPTVEVLVSGKWTPPRVRYSPFLPKE 803
Db SAVERDLVVKGLSYGTGTAVESLOASRPILPTVEVLVSGKWTPPRVRYSPFLPKE 845
QY 804 PREDKSHHPKDRGRFVSLHNSVLSNQVEQPDPRPARVAGSWGPCSCSGGLQKRA 863
Db PREDKSHHPKDRGRFVSLHNSVLSNQVEQPDPRPARVAGSWGPCSCSGGLQKRA 878
QY 846 ----RES-----FNAIPTFSS-----AWVTEWGECSKCELGWQRRLL 878

QY 864 VDCRSAGQRTVPACDAAH---RPVETOACGE-PCPTWELSAWSPCKSCGQGRFORSIKC 920
Db VECEDINGQ---PASECAKEVKPASTRPCADHCFQPOLGWSCKSCGKGYKKRSIKC 935
QY 921 VGHGRLRLAQCNLHHRKQOE-LDFCVLRPC 950
Db LSHDGGVLSHESCDPLKKPKRHFDICTMARC 966
RESULT 4
US-09-321-987B-4
; Sequence 4, Application US/09321987B
; Patent No. US20020102210A1
; GENERAL INFORMATION:
; APPLICANT: Kimble, Judith E
; APPLICANT: Bielloch, Robert H
; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
; FILE REFERENCE: 960296.95386
; CURRENT APPLICATION NUMBER: US/09/321,987B
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,170
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/129,023
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Murine
US-09-321-987B-4

Query Match 48.1%; Score 2485.5; DB 10; Length 950;
Best Local Similarity 48.0%; Pred. No. 5.2e-169;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
QY 1 MLLGILTLAFAGTAGG---SEPEREVVPIRLDPDINGRRYVRGP-EDSGQGLIFQI 57
Db LLLASITMLLCARGHGRTEDEELVLP-SLE-----RAPCHDSTTRL--RL 66
QY 58 TAFQEDYLLHTLPDQAFLAPAFSTHGLGVPLQGLTGS-----SDLRRCFYSGDVNA 109
Db DAFQQLHLKLPDQSGFLAPGFTLQTV-----GRSPGSEAQLHDPDGLAHCEYSGTVNG 121
QY 110 EPDSFAVSLCGGLRGAFYGAAYVISPPLNAS---APAAQNSOGA-----HLLQRRG 161
Db DPGSAALSLCEGVRGAFYLOGEEFFIQAPAGVATERLAPAVPEESSARPQHILRRR- 180
QY 162 VPGGPGDPTSRCSGVASGWNPAILRALDPYKPRRAGFSGSRSSRRR----- 206
Db RRGSGGAKCGVMD-----DELPTSDSPESQNTNQPVRDPTPDAGKP 226
QY 207 ----SGRAKRFVSPRYVETLVVADESVMKFGHADLEHYLLTLATAARLYRHPSILNFI 262
Db SGPSIRKRRFVSPRYVETMLVADQSMADFHGSLKHYLLTLFVAARLYKHPSIRNSI 286
QY 263 NIWVVKVLLLRDRSGKYVGNALTLRNCANQKLNKYSDKHPEYWDTAIFYHQDL 322
Db SLVVVKVLTVEEQKGPVTSNAALTLRNCANQKLNKYSDKHPEYWDTAIFYHQDL 346
QY 323 GATTCDTLGMADVGTMDCKPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVVCVEFG 382
Db GATTCDTLGMADVGTMDCKPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVVCVEFG 406
QY 383 KLRANHMSPFLTIQIDRANPWSACSAIITDLDGSHGDCLLDQPKSPISLPEDLPASV 442
Db KLRANHMSPFLTIQIDRANPWSACSAIITDLDGSHGDCLLDQPKSPISLPEDLPASV 466
QY 443 TLSQOCELAFGVGSKPCP-YWQYCTKLWCTGKAKQWVCOTRHPFADGTSCGEGKLC 501
Db DANRQCOFTFGESKHCPCDAASTCTLWCTGSGGLVLCQTKHFPWADGTSCGEGK 526
QY 502 GACVVERHNLNKH---RVDGSKWAKWDPYGSCRTCGGGVQLARRQCTNPTPANGSKYCEGV 558

Db 527 GKCVNKTDM-KHEATPVHSGWPGWGDSCRTCGGVYTHRECDNPVPKNGKCYCBK 585
QY 559 RVKYSNLEPCPSSASGSKFREEOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRCKKL 618
Db 586 RVKYSNLEPCPSSASGSKFREEOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRCKKL 644
QY 619 ICRANGTGYFVVLAPKVVVDTGLCSPTDSTVSCVQGVKAGCDNLGSKKRFKCGVCGGD 678
Db 645 TCEAKGIGYFVVLQPKVVVDTGLCSPTDSTVSCVQGVKAGCDNLGSKKRFKCGVCGGD 704
QY 679 NKSCKVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLICDDNYLALKNSQGYLLN 738
Db 705 GSTCKKMSGIVSTRPGYHDIVTPAGATNIEVKHNRQGRNNGSFLAIRAADGTIYL 764
QY 739 GHEVVAVERDLVYKGLSLRYSGTGTAVESLQASRIEPLTVEVLSVGKMTPPRVRYSF 798
Db 765 GNFTLSTLEQDLYKTGLVRYSGSSAALERSFSPLKEPLTIQVLMVGHALRPKIKETY 824
QY 799 YLPKPREDKSSHPKDPGRGSPVLHNSVLSNOVEQDPPRPAWVAGSGPCSCSGSG 858
Db 825 FMKKKTES-----FNAIPTFS-----FWVIEENGECSTKCGSG 857
QY 859 LOKRAVDCRGSGAGRTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSCGRGFOR 915
Db 858 WQRRVVOQCRDINGH---PASECAKEVPKASTRCPADLPCPHQVGDWSPCKTCGKYKK 914
QY 916 RSLKCVGHGGRLLARDQCNLHRRKQPE-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCDPLKPKHYIDFCTLTQC 950

RESULT 5
US-10-163-316-7
; Sequence 7, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI01-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-163-316-7

Query Match 48.1%; Score 2485.5; DB 9; Length 968;
Best Local Similarity 48.0%; Pred. No. 5,4e-169;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

QY 1 MLLGLTILAFAGRTAGG--SPEREVVVPIRLDPDINGRRYWRGP-EDSGDQGLIFQI 57
Db 37 LLLLASTTLLCARGAHRPTDEELVLP-SLE-----RAPGHDSTTTL--RL 83
QY 58 TAPQEDFYHLHTPDQAFAPASTEHLGVPLQGLTGG-----SDLRRCFYSGDVNA 109
Db 84 DAFGQQLHLKLOPDSGLFAPGTTLQTV-----GRSPGSEAOHLDPDGLAHCFYSGTVNG 138
QY 110 EPDSFAVSLCGLRGAFYGAEEYVISPLPNAS-----APAAQRNSQGA-----HLLQRRG 161
Db 139 DPGSAALSCLCEGVRCFAYLQGEFFIQPAGVATERLAPVPEESSARPQFHLIRRR- 197
QY 162 VPGGPGSDPTSCGVASGNWPAIRALRDYPKPRRAGFCESRRRR----- 206
Db 198 -----RRSGGAKCGVMD-----DETLPSTSDRPESQNTNQWVPVRDPTQDAGKP 243

QY 207 ----SGRAKRFVSTPRVYETLVVADESNKVFHGADLBHYLLTLLATAARLYRHPSTINPI 262
Db 244 SGPGSKRRKREPRVYETMLVADQSMADFHGSLGLHYLLTFSVAAREYKHPSPNSI 303
QY 263 NIVVKKVLLLRDSDGPKVTGNAALTIRNFCAMOKKLINKYSDKHPEYWDTAILFTTRCLC 322
Db 304 SLVVKKLVIIYEEOKGPEVTSNAALTIRNFCMOKQINSFSDRDPEDHYDTAILFTTRDLC 363
QY 323 GATTCDTLGMADVMTGDKPRKSCSVIEDDGLPSAFTTAHELGHVFNPMHDPNVKVCVEVFG 382
Db 364 GSHTCDTLGMADVMTGDKPRKSCSVIEDDGLPSAFTTAHELGHVFNPMHDPNVKVCVEVFG 423
QY 383 KLRANHMSPPTLIODIRANPWSAGSAAIITDFDSHGDCLLDQPSKPISELPEDLPGASY 442
Db 424 VTGSHLMASMLSLDHSQWSPSCSAYMVTFSFLDNHGHECLMDKPNQPIKLPDLPGLTY 483
QY 443 TLSOCELAFCGVGSKPCP-YMYCTCKLWCTGKAGOMVOCOTRHFPMADGTSCGEGKLCCK 501
Db 484 DANQCOPTFGESKHCPCDAASTCTTLWCTGTSGLLVLCQTKHPPWADGTSCGEGKWCVS 543
QY 502 GACVYERHNLANKH---RVDSGWAKWDYPCGSRCTCGGVQLARRQCTNPTPANGSKYCEV 558
Db 544 GKCVNKTDM-KHEATPVHSGWPGWGDSCRTCGGVYTHRECDNPVPKNGKCYCBK 602
QY 559 RVKYSNLEPCPSSASGSKFREEOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRCKKL 618
Db 603 RVKYSNLEPCPSSASGSKFREEOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRCKKL 661
QY 619 ICRANGTGYFVVLAPKVVVDTGLCSPTDSTVSCVQGVKAGCDNLGSKKRFKCGVCGGD 678
Db 662 TCEAKGIGYFVVLQPKVVVDTGLCSPTDSTVSCVQGVKAGCDNLGSKKRFKCGVCGGD 721
QY 679 NKSCKVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLICDDNYLALKNSQGYLLN 738
Db 722 GSTCKKMSGIVSTRPGYHDIVTPAGATNIEVKHNRQGRNNGSFLAIRAADGTIYL 781
QY 739 GHEVVAVERDLVYKGLSLRYSGTGTAVESLQASRIEPLTVEVLSVGKMTPPRVRYSF 798
Db 782 GNFTLSTLEQDLYKTGLVRYSGSSAALERSFSPLKEPLTIQVLMVGHALRPKIKETY 841
QY 799 YLPKPREDKSSHPKDPGRGSPVLHNSVLSNOVEQDPPRPAWVAGSGPCSCSGSG 858
Db 842 FMKKKTES-----FNAIPTFS-----FWVIEENGECSTKCGSG 874
QY 859 LOKRAVDCRGSGAGRTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSCGRGFOR 915
Db 875 WQRRVVOQCRDINGH---PASECAKEVPKASTRCPADLPCPHQVGDWSPCKTCGKYKK 931
QY 916 RSLKCVGHGGRLLARDQCNLHRRKQPE-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDPLKPKHYIDFCTLTQC 967

RESULT 6
US-10-097-597-1
; Sequence 1, Application US/10097597
; Publication No. US20030022352A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiichi
; APPLICANT: Hakezaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same
; TITLE OF INVENTION: pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADA
; FILE REFERENCE: 057092
; CURRENT APPLICATION NUMBER: US/10/097,597
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03

; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-597-1

Query Match 44.1%; Score 2274; DB 9; Length 727;
Best Local Similarity 54.4%; Pred. No. 4.5e-154;
Matches 408; Conservative 129; Mismatches 173; Indels 40; Gaps 11;
QY 209 RAKRFVSIPRYVETLVVADESVMVKGADLEHYLLTLLATAARLYRHPHILNPINIVVVK 268
Db 9 RKKRFVSSPRYVETLVVADESVMVKGADLEHYLLTLLATAARLYRHPHILNPINIVVVK 68
QY 269 VLLLRDRSGPKVTGNAALTLRNFCAMQKLNKYSKHPKPEYWDTAIFTRQDLGGATCD 328
Db 69 ILVIHQEGPEVTSNAALTLRNFCAMQKLNKYSKHPKPEYWDTAIFTRQDLGGATCD 128
QY 329 TLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNDNVKCVLFGKLRANH 388
Db 129 TLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNDNVKCVLFGKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAAITDLDLSGHGCLLDQPSKIPSLPEDLFCASVTLSSQC 448
Db 189 MMSPTLIQIDRANPWSACSAAITDLDLSGHGCLLDQPSKIPSLPEDLFCASVTLSSQC 248
QY 449 ELAFGVSGKPCP-YMQYCTKLWCTGKAKQMVCTRHFPWADGTSCEGKCLCKGACVER 507
Db 249 QFTFGEDSKHCPDAASTCTLWCTGSGVLVCQTKHFPWADGTSCEGKCLCKGACVER 308
QY 508 HNLNKH---RVDGSKAKWDPYPCSRCTCGGVQLARRQCTNPTPANGKCYCEGVRYKRS 564
Db 308 TD-RKHFDTPFHGSMGPGWPCDSCRTCGGVQYTMRECDNPVKNGKCYCEGVRYKRS 367
QY 565 CNLEPCPSASGKSFREEOCEAHNFPSKASFGSGPAVEWIPKYAGVSPKRCCLICRANG 624
Db 368 CNLEPCPSASGKSFREEOCEAHNFPSKASFGSGPAVEWIPKYAGVSPKRCCLICRANG 426
QY 625 TGYFVFLAPKVVDGTLCSPDSTSVCGQKCIKAGCDNGLSKRFRDCKGCGGDNKSKCK 684
Db 427 TGYFVFLAPKVVDGTLCSPDSTSVCGQKCIKAGCDNGLSKRFRDCKGCGGDNKSKCK 486
QY 685 VTGLFTKPMHGYNFWAIPAGASSIDIRQYKGLIGDDNYLALNKSQKYLNGHFVVS 744
Db 487 ISGSVTSAPKGYHDIVTIPTGATNIEVKQNRQSRNNGSFLAIAADGTYILNGDYTL 546
QY 745 AVERDLVVKGLLRYSGTGTAVESLOASRPILPTVEVLSVGKMTPPRVRYSFYLPEP 804
Db 606 ---KES-----FNAITFS-----AWIEEGECSKCELGWQRRLV 639
QY 865 DCRGSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSGCFQRRSLKCV 921
Db 640 ECRDINGQ--PASECAKEVKPASTRCPADHPCPQWOLGEMSSCKTCKGKYKRSKLCL 696
QY 922 GHGRLIARDQCNLHRKPEQ-LDFCVLRPC 950
Db 697 SHDGVLSHESCDPLKPKHFIDFCTLTQC 726

RESULT 7
US-10-097-580-1
; Sequence 1, Application US/10097580
; Publication No. US20030032168A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka

; APPLICANT: Inouchi, Ei-ji
; APPLICANT: Hakezaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/10/097,580
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-580-1
Query Match 44.1%; Score 2274; DB 9; Length 727;
Best Local Similarity 54.4%; Pred. No. 4.5e-154;
Matches 408; Conservative 129; Mismatches 173; Indels 40; Gaps 11;
QY 209 RAKRFVSIPRYVETLVVADESVMVKGADLEHYLLTLLATAARLYRHPHILNPINIVVVK 268
Db 9 RKKRFVSSPRYVETLVVADESVMVKGADLEHYLLTLLATAARLYRHPHILNPINIVVVK 68
QY 269 VLLLRDRSGPKVTGNAALTLRNFCAMQKLNKYSKHPKPEYWDTAIFTRQDLGGATCD 328
Db 69 ILVIHQEGPEVTSNAALTLRNFCAMQKLNKYSKHPKPEYWDTAIFTRQDLGGATCD 128
QY 329 TLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNDNVKCVLFGKLRANH 388
Db 129 TLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNDNVKCVLFGKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAAITDLDLSGHGCLLDQPSKIPSLPEDLFCASVTLSSQC 448
Db 189 MMSPTLIQIDRANPWSACSAAITDLDLSGHGCLLDQPSKIPSLPEDLFCASVTLSSQC 248
QY 449 ELAFGVSGKPCP-YMQYCTKLWCTGKAKQMVCTRHFPWADGTSCEGKCLCKGACVER 507
Db 249 QFTFGEDSKHCPDAASTCTLWCTGSGVLVCQTKHFPWADGTSCEGKCLCKGACVER 308
QY 508 HNLNKH---RVDGSKAKWDPYPCSRCTCGGVQLARRQCTNPTPANGKCYCEGVRYKRS 564
Db 308 TD-RKHFDTPFHGSMGPGWPCDSCRTCGGVQYTMRECDNPVKNGKCYCEGVRYKRS 367
QY 565 CNLEPCPSASGKSFREEOCEAHNFPSKASFGSGPAVEWIPKYAGVSPKRCCLICRANG 624
Db 368 CNLEPCPSASGKSFREEOCEAHNFPSKASFGSGPAVEWIPKYAGVSPKRCCLICRANG 426
QY 625 TGYFVFLAPKVVDGTLCSPDSTSVCGQKCIKAGCDNGLSKRFRDCKGCGGDNKSKCK 684
Db 427 TGYFVFLAPKVVDGTLCSPDSTSVCGQKCIKAGCDNGLSKRFRDCKGCGGDNKSKCK 486
QY 685 VTGLFTKPMHGYNFWAIPAGASSIDIRQYKGLIGDDNYLALNKSQKYLNGHFVVS 744
Db 487 ISGSVTSAPKGYHDIVTIPTGATNIEVKQNRQSRNNGSFLAIAADGTYILNGDYTL 546
QY 745 AVERDLVVKGLLRYSGTGTAVESLOASRPILPTVEVLSVGKMTPPRVRYSFYLPEP 804
Db 547 TLEQDIMYKGVVLRYSGSSAALRISFSLKEPLTIQVLTGNALRPKIYTFYVKK- 605
QY 805 REDKSSHPKDRGSPSVLHNSVLSLNSQVBPDDPPARWVAGSPGCSASCGSLQKRAV 864
Db 606 ---KES-----FNAITFS-----AWIEEGECSKCELGHQRRLV 639
QY 865 DCRGSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSGCFQRRSLKCV 921
Db 640 ECRDINGQ--PASECAKEVKPASTRCPADHPCPQWOLGEMSSCKTCKGKYKRSKLCL 696
QY 922 GHGRLIARDQCNLHRKPEQ-LDFCVLRPC 950
Db 697 SHDGVLSHESCDPLKPKHFIDFCTLTQC 726

```
Db 640 ECRDINGQ---PASECAKEVKPASTRPCADHPCCPQWQLGEMSSCKTCGKGYKKRSLKCL 696
QY 922 GHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 697 SHDGGVLSHESCDPLKKPKHFDICTLTQC 726

RESULT 8
US-09-445-023A-1
; Sequence 1, Application US/09445023A
; Patent No. US20020119167A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-445-023A-1

Query Match 44.1%; Score 2274; DB 10; Length 727;
Best Local Similarity 54.4%; Pred. No. 4,5e-154;
Matches 408; Conservative 129; Mismatches 175; Indels 40; Gaps 11;

QY 209 RAKRFVSPRYVETLVVADSMVKFHGADLEHYLLTLATAARLYRHPSILNFINIVVK 268
Db 9 RKKRFVSSPRYVETLVVADSMVAFHSGGLKHYLLTLFSAARLYKHPSIRKNSVSLVVK 68
QY 269 VLLLRDRSDGPKVTGNAALTLRNFCAWOKKLNKVSDDKHPYWDTAILFTFQDLGGATCD 328
Db 69 ILVIHEQKGPEVTSNAALTLRNFCAWOKKLNKVSDDKHPYWDTAILFTFQDLGGATCD 128
QY 329 TLGMADVTGTCMDPKRSCVIEDDGLPSAFTTAHELGHVENMHPDNVKNVCEVEFGKLRANH 388
Db 129 TLGMADVTGTCMDPKRSCVIEDDGLPSAFTTAHELGHVENMHPDNVKNVCEVEFGKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAALITDLDSDHGDCLLDQPSKPISLPEDLPFASVYLSQOC 448
Db 189 MMSPTLIQIDRANPWSACSAALITDLDSDHGDCLLDQPSKPISLPEDLPFASVYLSQOC 248
QY 449 ELAFGVSGKPCP-YMQYCTKLWCTGKAGQMVQCTRHFPWADGTSCGEGKILCKGACVER 507
Db 249 QFTFGEDSKHCPDAASTCTTLWCTGSGGLVLCQTKHFPWADGTSCGEGKILCKGACVER 308
QY 508 HNLNKH---RVDGSAWAKWDPYGCSTCGGVQLARRQCTNPTPANGKCYGEGVRVYKRS 564
Db 309 TD-RKHEDTPFHSGSWGPGWDCSTRTCGGVQVYTHRECDNVPKNGKCYGEGVRVYKRS 367
QY 565 CNLEPCSSASGKSFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPKCKLICRANG 624
Db 368 CNLEDCPDN-NGKTFREOCEAHNEFSKASFGSGPAVEIPKYAGVSPKDRCKLICQAKG 426
QY 625 TGFFVVLAPKVVDGTLCSPDSTSVQVQKCIKAGDGNLGSKKRPDKCVCVGGDNKSKCK 684
Db 427 IGYFFVLQPKVDGTPCSPDSTSVQVQKCIKAGDGNLGSKKRPDKCVCVGGDNKSKCK 486
QY 685 VTGLFKPKHGFVNFVAIPAGASSIDIRQYKGLITGDDNLYLALNSQCKYLLNGHFVYS 744
Db 487 ISGSVTSARPGYHDIIVITGATNIEVKQNRQGRNNGSFLAIKAADGTIILNGDYTLIS 546
```

```
QY 745 AVERDLVVGSLRLRYSGTGTAVESLQASRPILPLVEVLSVGKMTTPRVRYSFYLPKEP 804
Db 547 TLEQDIMYKGVLRYSGSSAALERISFSLKEPLTFTQVLTGVNLRPKIKYIFYVKKK- 605
QY 805 REDKSSHPKDPGRGSPVLSVLSNQVEQDDPRPARVWAGSWGSCSACSGSLQKRAV 864
Db 606 ---KES-----FNAITFS-----AWVIEEGECSCSELGWRRLV 639
QY 865 DCRGSAGORTVPACDAH--RPVETOACGE-PCPTWELSAWSPCKSGRFGORRSLKCV 921
Db 640 ECRDINGQ---PASECAKEVKPASTRPCADHPCCPQWQLGEMSSCKTCGKGYKKRSLKCL 696
QY 922 GHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 697 SHDGGVLSHESCDPLKKPKHFDICTLTQC 726

RESULT 9
US-10-097-597-12
; Sequence 12, Application US/10097597
; Publication No. US20030022352A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,
; TITLE OF INVENTION: pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/10/097,597
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-097-597-12

Query Match 44.0%; Score 2273; DB 9; Length 727;
Best Local Similarity 53.7%; Pred. No. 5.4e-154;
Matches 403; Conservative 131; Mismatches 176; Indels 40; Gaps 10;

QY 209 RAKRFVSPRYVETLVVADSMVKFHGADLEHYLLTLATAARLYRHPSILNFINIVVK 268
Db 9 RKKRFVSSPRYVETLVVADSMVAFHSGGLKHYLLTLFSAARLYKHPSIRKNSVSLVVK 68
QY 269 VLLLRDRSDGPKVTGNAALTLRNFCAWOKKLNKVSDDKHPYWDTAILFTFQDLGGATCD 328
Db 69 ILVIHEQKGPEVTSNAALTLRNFCAWOKKLNKVSDDKHPYWDTAILFTFQDLGGATCD 128
QY 329 TLGMADVTGTCMDPKRSCVIEDDGLPSAFTTAHELGHVENMHPDNVKNVCEVEFGKLRANH 388
Db 129 TLGMADVTGTCMDPKRSCVIEDDGLPSAFTTAHELGHVENMHPDNVKNVCEVEFGKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAALITDLDSDHGDCLLDQPSKPISLPEDLPFASVYLSQOC 448
Db 189 MMSPTLIQIDRANPWSACSAALITDLDSDHGDCLLDQPSKPISLPEDLPFASVYLSQOC 248
QY 449 ELAFGVSGKPCP-YMQYCTKLWCTGKAGQMVQCTRHFPWADGTSCGEGKILCKGACVER 507
Db 249 QFTFGEDSKHCPDAASTCTTLWCTGSGGLVLCQTKHFPWADGTSCGEGKILCKGACVER 308
QY 508 HNLNKH---RVDGSAWAKWDPYGCSTCGGVQLARRQCTNPTPANGKCYGEGVRVYKRS 564
```


[illegible]

```

329 QY TLGMADVTCMDCKPRSCSVIEDDGLPSAFTTAHELGHVFNPNHDNVKCEVEYFGKLRANI 388
Db TLGMADVTCMDCKPRSCSVIEDDGLPSAFTTAHELGHVFNPNHDNVKCEVEYFGKLRANI 389
129 QY TLGMADVTCMDCKPRSCSVIEDDGLPSAFTTAHELGHVFNPNHDNVKCEVEYFGKLRANI 390
Db TLGMADVTCMDCKPRSCSVIEDDGLPSAFTTAHELGHVFNPNHDNVKCEVEYFGKLRANI 391
389 QY MMSPTLIOIDRANPWSACSAIITDFLDSHGDCLLDQSKPISLPELPGASYTLSSQOC 448
Db MMSPTLIOIDRANPWSACSAIITDFLDSHGDCLLDQSKPISLPELPGASYTLSSQOC 449
189 QY LMAWMLSELDSHQWSPCSAYWVTSFLDNGHGECMLDKPQNIKLPSDLPTLYLDNRQC 248
Db LMAWMLSELDSHQWSPCSAYWVTSFLDNGHGECMLDKPQNIKLPSDLPTLYLDNRQC 249
449 QY ELAFEGVSKPCP-YMQYCTKLMCTCKAKGQWVCQTRHFPWADGTSCEGKLCCLKGACVER 507
Db ELAFEGVSKPCP-YMQYCTKLMCTCKAKGQWVCQTRHFPWADGTSCEGKLCCLKGACVER 508
249 QY QFTFGEESKPCDAASTCTTLMCTGTSGGLLVQCQTHFPWADGTSCEGKWCYSGKCVNK 308
Db QFTFGEESKPCDAASTCTTLMCTGTSGGLLVQCQTHFPWADGTSCEGKWCYSGKCVNK 309
508 QY HNLANKH---RVDGSAWKWDVPCPSRRTCGGVQLARQCTNPTPANGKKYCEGVRVKYS 564
Db HNLANKH---RVDGSAWKWDVPCPSRRTCGGVQLARQCTNPTPANGKKYCEGVRVKYS 565
309 QY TDM-KHFATPVHGSWGPWGDSCRTCGGVQYTMRECDNPVKNGKCYCEGKRVYRS 367
Db TDM-KHFATPVHGSWGPWGDSCRTCGGVQYTMRECDNPVKNGKCYCEGKRVYRS 368
565 QY CNLEPCSSASGSKFREQEAFNGYNHSTNRLTLAVAWVPKYSVSPDRCKCLICRANG 624
Db CNLEPCSSASGSKFREQEAFNGYNHSTNRLTLAVAWVPKYSVSPDRCKCLICRANG 625
368 QY CNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKAGVSPDRCKLTCEAG 426
Db CNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKAGVSPDRCKLTCEAG 427
625 QY TGYFVVLAPKVVDGTLCSPOSTSVCVQOGKIKAGCDGNLCKRRPKDCKGVCGDGNKSCKK 684
Db TGYFVVLAPKVVDGTLCSPOSTSVCVQOGKIKAGCDGNLCKRRPKDCKGVCGDGNKSCKK 685
427 QY IGYFVVLQPKVVDGTLCSPOSTSVCVQOGCVKAGCDRIIDSKKKFKDCKGVCGDGNKSCKK 486
Db IGYFVVLQPKVVDGTLCSPOSTSVCVQOGCVKAGCDRIIDSKKKFKDCKGVCGDGNKSCKK 487
685 QY VTGLFTPMHGYFVWAI PAGASSIDIRQRYGKGLIGDDNYLALKNSOGKYLNLGHVVS 744
Db VTGLFTPMHGYFVWAI PAGASSIDIRQRYGKGLIGDDNYLALKNSOGKYLNLGHVVS 745
487 QY MSGIVTSTRPYHDIVIPAGATNIEVKHRNQGRNNGSFLAIRAADGTYILNGFTLS 546
Db MSGIVTSTRPYHDIVIPAGATNIEVKHRNQGRNNGSFLAIRAADGTYILNGFTLS 547
745 QY AVERDLVVKGLLRYSGTGAVESLQASRPILPEPLTVEVLSVGKMTTPRVYSFYLPKEP 804
Db AVERDLVVKGLLRYSGTGAVESLQASRPILPEPLTVEVLSVGKMTTPRVYSFYLPKEP 805
547 QY TLEQDLVTKGTVLRYSGSSAALERSFSLKEPLTQVLMVGHALRPKIKFTYFMKKYT 606
Db TLEQDLVTKGTVLRYSGSSAALERSFSLKEPLTQVLMVGHALRPKIKFTYFMKKYT 607
805 QY REDKSSHPKDPGRGSLVHNSVLSISNOVEQDDPRPARVAGSWGSPCSASCSGLQKRAY 864
Db REDKSSHPKDPGRGSLVHNSVLSISNOVEQDDPRPARVAGSWGSPCSASCSGLQKRAY 865
607 QY ES-----FNAIPYFS-----EWIEEWGBCSKTCSGSGWQRRV 639
Db ES-----FNAIPYFS-----EWIEEWGBCSKTCSGSGWQRRV 640
865 QY DCRGSAGQRTVPACDAAH--RPVETQAGE-PCPTWELSAWSPCSKSGRQFRRSLKCV 921
Db DCRGSAGQRTVPACDAAH--RPVETQAGE-PCPTWELSAWSPCSKSGRQFRRSLKCV 922
640 QY QCRDINGH---PASECAKEVKPASTRCPADLPCHQVGDWSPCSKTCGKGYKRTLKCV 696
Db QCRDINGH---PASECAKEVKPASTRCPADLPCHQVGDWSPCSKTCGKGYKRTLKCV 697
922 QY GHGGRLLARQCNLHRKPQE-LDFCVLRPC 950
Db GHGGRLLARQCNLHRKPQE-LDFCVLRPC 951
697 QY SHDGGVLSNESCDDLKPKPHYIDFCTLQTC 726
Db SHDGGVLSNESCDDLKPKPHYIDFCTLQTC 727

RESULT 11
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. US20020119167A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiichi
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical composition and method of immunologically analyzing
; FILE REFERENCE: C57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

```

Query Match 44.0%; Score 2273; DB 10; Length 727;

[illegible]

QY 361 HELG 364
Db 361 HELG 364

RESULT 15

US-10-174-590-352
; Sequence 352, Application US/10174590
; Publication No. US2003008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 352
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-352

Query Match 37.1%; Score 1914; DB 9; Length 837;
Best Local Similarity 46.3%; Pred. No. 2.7e-128;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

QY 1 MLLGLTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGQLFIQTAF 60
Db 37 LLLLLLASLLPSARLASPLPREEEIVFEKLGSLV-----PGSAGAFARLLCKLQAF 88
QY 61 OEDFYHLHLPDAQFLAPASTHEHLVPLQGLTGGSSDLRRCFYSGDVNAEFDSPFAVSLC 120
Db 89 GETLLELEQDSGVQVQVEGLTVQYLGQAP-LLGGAEP--GTYLTGTINCEPESVASLHWD 145
QY 121 GG-LRGAFGYRGAEEYVISPLPNASAPAAQNSOGAHLLORRGVPGPGSGTTSRCGVASG 179
Db 146 GGALLGVLYRGALHLQPLEGGTPNSA--GGPGAHILRRK-----SPASSGGPMCNV--- 196
QY 180 WNPAILRALDPYKPRRAGGESRRSRGRKREVSIPRYVETLVVADSMVKFRGADLE 239
Db 197 -----KAPLGSPSPRPR--RAKRFASLSRFVETLVVADDKMAAFHGAGLK 239
QY 240 HYLLTLATAARYLRHPSILNINIVVKKVLLLRDRDGPVKTGNAALTLRNFCAMQKLL 299
Db 240 RYLLTYMAAAAKAFKPSIRPNVSLVTVLVLGSGEGEPQVGPAAQTLRSFCAMQRL 299
QY 300 NKVSDKHPEYWDFTAILFTQDLGATTCTGLGMADVMTCDPKRSCSVIEDDGLPSAFTT 359
Db 300 NTPEDSGPHDFTAILFTQDLGCVSTCDTLGMADVMTCDPARSCAIVEDDGLQSAFTA 359
QY 360 AHELGHVFNHNDNVKVEEVFKL-RANHMSPTLIQIDRANPWSACSAAIITDFLDSG 418
Db 360 AHELGHVFNHNDNSKPCISLNGPLSTSRHVAPMAHVDPPEEPWSPCSARFITDFLDNG 419
QY 419 HGDCILLDQSPKPSLPELPLGASYTLISQOCELAFGVGSKPCPYM-QYCTKLMCTCKAKGQ 477
Db 420 YGCHLLDKPEAPLHLVPTTGGDYDADQCQLTFGPDSDRHCPLPPPCALWCSGHLNGH 479
QY 478 MVCQTRHFWADGTCGEGKCLKLGACVERHNLNKHVD--GSAKWDPDYPGCSRFTCGGG 535
Db 480 AMCQTKHSWADGTCGPAQACMGGRCLHWDQLQDENIFQAGGWGPGWPGWDCSRTCGGG 539

QY 536 VQLARROCTNPTPANGKYCEGVYRVKYSNLEPCPSSASGKSFREEOCEAFNGYNHSTN 595
Db 540 VQFSSRDCTRPVPRNGKYCEGRRTFRSCNTEDCP-TGSALTFRREQCAA---YNHRTD 595
QY 596 RLTL---AVAWPKYSGVSPRDCKKLCICRANGTGYFYVLAPKVVVDGTLCSPDSTSYCVQ 652
Db 596 LFKSFGPMQMDVPRYTGVAPDQCKLTQCARALGYTVVLEPRYVDGTPCSPDSSSVCVQ 655
QY 653 KCIKAGCDNGLGSKKRFDKGCGGDNKSCCKVTGLTKPMHGYNFVVAIPAGASSIDIR 712
Db 656 KCHAGCDRIIGSKKFKDKCMVCGGSGSGKSGGFRKFRYGINNVVITIPAGATHILVR 715
QY 713 QRGYKGLIGDDNYLALKNSOGKYLLNGHFVVSAYERDLVVKSL-LRYSGTGTAVESLCA 771
Db 716 QOQNGP--HRSIYLALKLPDGSYALNGEYTLMPSPDVLPGAVSLRYSGATAASETLSG 773
QY 772 SRPILEPLTVEVLSVGKMTPPRVRYSPYLPKEPREDKSSHPKD 814
Db 774 HGPLAQPLTLQVLVAGNPQDTRLRYSEFFVPRPTPSTPRPTPD 816

Search completed: May 9, 2003, 15:30:18
Job time : 63 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 18:46:45 ; Search time 3590 Seconds
(without alignments)
12870.675 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgctctgctgagcatcct.....gcgtctgagcgctgctga 2853

Scoring table:
IDENTITY: NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estchum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query Match	Length	DB ID	Description
1	632.4	22.2	2931	11 BC003269	BC003269 Mus muscu
2	530.2	18.6	849	13 B1103177	B1103177 602889417
3	511	17.9	681	13 BG921734	BG921734 602825569
4	409.2	14.3	961	10 BB612189	BB612189 BB612189
5	391.4	13.7	522	10 BE553572	BE553572 ut47B04.Y
6	345.6	12.3	544	12 BF078689	BF078689 229090 MA

7	339.4	11.9	627	10 BE281680	BE281680 601099657
8	270.6	9.5	654	10 BE016461	BE016461 F733410.X
9	263.4	9.2	738	13 B1819909	B1819909 603035454
10	254.8	8.9	746	13 B1733795	B1733795 603351940
11	245.8	8.6	661	10 BB220373	BB220373 BB220373
12	244.2	8.6	660	10 BB651049	BB651049 BB651049
13	242	8.5	662	10 BB642806	BB642806 BB642806
14	240.2	8.4	951	14 BQ930852	BQ930852 AGENCOURT
15	229.8	8.1	945	12 BF033017	BF033017 60145701
16	227.8	8.0	945	12 BF033017	BF033017 60145701
17	226	7.9	984	17 BMS70576	BMS70576 1103910.Y
18	224.4	7.9	797	14 BQ58765	BQ58765 AGENCOURT
19	222.6	7.8	305	12 BF320986	BF320986 uz57H10.Y
20	219.6	7.7	598	13 BM069515	BM069515 1e8R09.X
21	215	7.5	674	10 BM225365	BM225365 BB225365
22	209.4	7.3	723	12 BG826611	BG826611 602748913
23	209.2	7.3	949	14 BQ893675	BQ893675 AGENCOURT
24	206.2	7.2	652	9 A1400905	A1400905 th30908.X
25	206	7.2	420	10 AW353152	AW353152 34940.MAR
26	205	7.2	681	12 BE912007	BE912007 601662774
27	202.8	7.1	562	13 BJ009737	BJ009737 BJ009737
28	201.2	7.1	602	10 AM137467	AM137467 UI-H-B11-
29	198.6	7.0	601	14 BQ332355	BQ332355 PM3-ET026
30	196.4	6.9	680	12 BF446232	BF446232 7p34C05.X
31	195.6	6.9	528	9 AA288689	AA288689 mp17f10.F
32	193.2	6.8	518	17 AZ374992	AZ374992 1M0128G10
33	188	6.6	770	17 CANS03W7R	CANS03W7R
34	186.4	6.5	689	12 BG242316	BG242316 602354488
35	185.6	6.5	667	10 BB624639	BB624639 BB624639
36	185.4	6.5	522	10 AW046509	AW046509 UI-M-BH1-
37	184.8	6.5	640	12 BE742606	BE742606 601575415
38	183.8	6.4	504	12 BG553803	BG553803 dac28a08.
39	181.6	6.4	543	14 BQ011430	BQ011430 UI-1-BCLP
40	180.6	6.3	468	10 BE664878	BE664878 UI-M-BH1-
41	179	6.3	475	9 A1148739	A1148739 qc65d04.X
42	178.4	6.3	264	10 BB576723	BB576723 BB576723
43	177.4	6.2	566	10 AV617111	AV617111 AV617111
44	174.6	6.1	257	9 A1644268	A1644268 vv85f01.X
45	172.4	6.0	464	10 BE666088	BE666088 149119 MA

ALIGNMENTS

RESULT 1
BC003269
LOCUS
DEFINITION
Mus musculus, Similar to a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 1, clone IMAGE:3499206, mRNA.
ACCESSION
BC003269
VERSION
BC003269.1 GI:14707747
KEYWORDS
HTC.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2931)
Srausberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT
NTH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-MGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amebcm.tmc.edu
 Gunatane, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. consortium/ILN: at: <http://image.llnl.gov>
 Series: IRK Plate: 9 Row: b Column: 7
 This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

source
 1. 2931
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="IMAGE:349206"
 /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI_CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector::pcmv-sport6"
 BASE COUNT 708 a 812 c 811 g 600 t
 ORIGIN

Query Match 22.2%; Score 632.4; DB 11; Length 2931;
 Best Local Similarity 61.0%; Pred. No. 1.5e-125;
 Matches 1083; Conservative 0; Mismatches 681; Indels 12; Gaps 3;

Db 631 AAGCGTTTCGTCTATCCCGGCTGAGTGGAGAGCTGTGTGCGGCGAGTCAATG 630
 1144 AAGGATTTGTGTCCACCCCGCTTATGTGMAACATGCTGCTGCGTGCACGATCATG 1203
 691 GTCAAGTTCCAGCGCGGAGCGAGACATTATCTGCTGACGCGTGGCGAGCGGCG 750
 1204 GCCACTTCCAGCGGAGCGGCTTAACGATTAACCTTCTCTGCTGCGGAGCG 1263
 751 CGACTACCGGCTCCAGCATCTCAACCCATCAACATCTGTGTGCAAGTGTG 810
 1264 AGGTTTAAAGCATCCAGCATAGAGATTAATTAATTAATTAATTAATTAATTA 1323
 811 CTCTTATGAGATGCTGATCCGCGGCGGAGGTCACGCGGAGTGGCGCTGAGGCTGCG 870
 1324 GTCAATATGAG 1383
 871 AACTTCTGTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 930
 1384 AATTTCTGCACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1443
 931 GACACTGCACTCTCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
 1444 GACACTGCAATTTCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1503
 991 GGCATGGGTATGTTGGTACCATGTGACCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1050
 1504 GGGATGGAGATGTTGGAGATGATGTGACCCAGAGAGAGAGAGAGAGAGAGAGAT 1563
 1051 GATGGGCTTCATCAGCTTCCAGCATGCGCAGAGAGAGAGAGAGAGAGAGAGAG 1110
 1564 GATGGTTCGAAAGCGGCTTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1623
 1111 CATGACATGGAAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
 1624 CAGCATGATGCTAGCACTGTGCGAGCTGATGATGATGATGATGATGATGATGAT 1683
 1171 TCCCGAGCCCTCATCAGATGAGCCGTCGCAACCCCTGTCAGAGCTGATGTCGATC 1230
 1684 GCTTCAGTGTCTCCAGCTTATGACATAGCAGCCCTGCTGACCTTACATGCTATAG 1743
 1231 ATCAACGACTTCTGAG 1290
 1744 GTCAAGTCTTCTAGATTAATGAGACAGGAGAGATGTTGATGAGCAAGCCCAAGAT 1803
 1291 ATCTCCGCGCGAGAGATGTGCGGCGCGCAGCTACACCTGAGCCGACAGTGTGAG 1350

Db 1804 ATCAAGCTCCCTTGTGATCTTCCCGGTACCTGTGATACGATGCCAACCGCAGATGCACTT 1863
 1351 GCTTTGGCGGTGGCTCCAGAGCCCTGTCTTACATGACATGAC---TGACCAACACTGTGG 1407
 1864 ACATTGGAGAGAGATCCAGCAGCTGCTGATGACAGCCAGCAGATGATGATGATGATG 1923
 1408 TGCACCGGAG 1467
 1924 TGCATGCGACCTCCGCTGCTTACTGTGTGTGTCGCAACAAACATCTTCCCTGGGAGAT 1983
 1468 GGCACCACTGTGGAG 1527
 1984 GGCACCACTGTGGAG 2043
 1528 CTCACAGCAGCAG-----GATGATGATGATGATGATGATGATGATGATGATGATG 1581
 2044 ATGAAGCATTTTGTCTACTCTGTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2103
 1582 TGCCTGCGACATGTTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1641
 2104 TGCCTAAGACCTGT 2163
 1642 CCTGCCAAGGAG 1701
 2164 CCAAG 2223
 1702 GAGCCCTGCGGAG 1761
 2224 GAG 2280
 1762 AACGGCTACACAG 1821
 2281 AATGATTTTCCAAAGCTTCTTGTGAGATGAGCCAGCTGTGATGATGATGATGATGATG 2340
 1822 TCCGCGCTGTCTCCCGGAG 1881
 2341 GCGGCGCTGTCCCAAG 2400
 1882 TCTATGCTGAG 1941
 2401 TTTTTCGCTTACAGCCAG 2460
 1942 GTCTGTGTCAG 2001
 2461 GTCTGTGTCAG 2520
 2002 AGATTGCAAG 2061
 2521 AAGTTGATAG 2580
 2062 CTCTTCACCAAGCCAG 2121
 2581 ATAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
 2122 AGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2181
 2641 AACTTGAAGTGAAG 2700
 2182 CTGAAG 2241
 2701 ATTAG 2760
 2242 CGGAG 2301
 2761 CAAG 2820
 2302 AGCTGAG 2361
 2821 AGAATCCGAG 2880
 2362 AAGATGACAG 2397

```

Db      2881 CATGCTCTCCGACCAAAATTAATTCACCTACTT 2916

RESULT 2
LOCUS   B1103177
DEFINITION B1103177 849 bp mRNA linear EST 26-JUN-2001
          60289417F1 NCI_CGAP_K1d14 Mus musculus cDNA clone IMAGE:5044493
          5', mRNA sequence.
ACCESSION B1103177
VERSION   B1103177.1 GI:14554070
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS   1 (bases 1 to 849)
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: Jeffrey E. Green, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L14M1122 row: a column: 06
           High quality sequence stop: 743.
           Location/Qualifiers
             1..849
               /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone_image="5044493"
               /lab_host="NCI_CGAP_K1d14"
               /note="Organ: kidney; Vector: pCMV-Sport6; Site:1: NotI;
               Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
               Average insert size 1.75 kb. Constructed by Life
               Technologies. Note: this is a NCI_CGAP library. 1"

BASE COUNT  193 a 250 c 250 g 156 t

ORIGIN
Query Match      18.6%; Score 530.2; DB 13; length 849;
Best Local Similarity 85.0%; Pred. No. 1e-103;
Matches 698; conservative 0; Mismatches 113; Indels 10; Gaps 9;

QY  1189 ATCGACCGTGGCAACCCC-TGGTCAGCTGTGAGTGCATCATCAC-CCGACTTCTCTGG
Db      1189 ATCGACCGTGGCAACCCC-TGGTCAGCTGTGAGTGCATCATCAC-CCGACTTCTCTGG
QY  1247 ACAGGGGGGAGGAGTACGCTGCTGGAGCAACCCAGCAAGCCCATCTCCCTCCGAGG 1306
Db      1247 ACAGGGGGGAGGAGTACGCTGCTGGAGCAACCCAGCAAGCCCATCTCCCTCCGAGG 1306
QY  86 ATAGTGGGATGTGATGCTGCTCTGGAGCACCCAGCAAGCCCATCTGCTCTGAGG 145
Db      86 ATAGTGGGATGTGATGCTGCTCTGGAGCACCCAGCAAGCCCATCTGCTCTGAGG 145
QY  1307 ATCTGCGGGGCGCAGCTACACCTGTAGCAGCACTGCGAGCTGCTTTGGCGTGGCT 1366
Db      1307 ATCTGCGGGGCGCAGCTACACCTGTAGCAGCACTGCGAGCTGCTTTGGCGTGGCT 1366
QY  146 ACCCTGCGGGGCAACAGTACATTTGAGCAGCACTGCGAGCTTGGGGTGGCT 205
Db      146 ACCCTGCGGGGCAACAGTACATTTGAGCAGCACTGCGAGCTTGGGGTGGCT 205
QY  1367 CCAAGCCCTGTCTTACATGACAGTACTGCACCAAGCTGTGTGACCGGGGAAGGCCAAG 1426
Db      1367 CCAAGCCCTGTCTTACATGACAGTACTGCACCAAGCTGTGTGACCGGGGAAGGCCAAG 1426
QY  206 CTAACCCCTGGCCATATATGCTACTGTACAAACCTGTGGTGCCTGCAAGGCCAAGG 265
Db      206 CTAACCCCTGGCCATATATGCTACTGTACAAACCTGTGGTGCCTGCAAGGCCAAGG 265
QY  1427 GACAGATGGTGTGAGAGCCGCGCCTTCCCTGGCGAGTGGCAAGCTGTGGCGAGG 1486
Db      1427 GACAGATGGTGTGAGAGCCGCGCCTTCCCTGGCGAGTGGCAAGCTGTGGCGAGG 1486
QY  266 GGCAATATGTGTGCGACAGCTGCGCACTTCCCTGGCGAGTGGCAAGCTGTGGCGAGG 325
Db      266 GGCAATATGTGTGCGACAGCTGCGCACTTCCCTGGCGAGTGGCAAGCTGTGGCGAGG 325
QY  1487 GCAGAGCTGTGCTCAAAAGGGGCTGCTGTGAGAGACACAACTTCAACAAGCACAGAGGTGG 1546
Db      1487 GCAGAGCTGTGCTCAAAAGGGGCTGCTGTGAGAGACACAACTTCAACAAGCACAGAGGTGG 1546
QY  326 GCAAGTCTGCTCAAGGAGAGCTGCTGTGAGAGACACAACTTCAACAAGTACCGGGTGG 385
Db      326 GCAAGTCTGCTCAAGGAGAGCTGCTGTGAGAGACACAACTTCAACAAGTACCGGGTGG 385

FEATURES
source

QY  1547 ATGTTCTCGGGCCAAATGGAGTCCCTATGAGCCCTGCTGCGACAGATGTGGGGCG 1606
Db      1547 ATGTTCTCGGGCCAAATGGAGTCCCTATGAGCCCTGCTGCGACAGATGTGGGGCG 1606
QY  386 ACGGCTCTTGGGCCAAGTGGGAGGCCCTTAGCTTCTGCTCCGACCTGCGGGGGCG 445
Db      386 ACGGCTCTTGGGCCAAGTGGGAGGCCCTTAGCTTCTGCTCCGACCTGCGGGGGCG 445
QY  1607 TGCAGCTGGCCAGAGGAGCGATGACACCAACCCACCCC-TGCCAAGGGGGGCAAGTACTGC 1665
Db      1607 TGCAGCTGGCCAGAGGAGCGATGACACCAACCCACCCC-TGCCAAGGGGGGCAAGTACTGC 1665
QY  446 TGCAGCTGGCCGAGGAGCGATGACCAACCTTACCCATGCAACGGTGGMAATACTGC 505
Db      446 TGCAGCTGGCCGAGGAGCGATGACCAACCTTACCCATGCAACGGTGGMAATACTGC 505
QY  1666 GAGGAGTGAAGGTGAATACCAGTCCCTGCAATCTGAGAGCCCTCCAGCTAGCTCC 1725
Db      1666 GAGGAGTGAAGGTGAATACCAGTCCCTGCAATCTGAGAGCCCTCCAGCTAGCTCC 1725
QY  506 GAGGAGTGAAGGTGAATACCAGTCCCTGCAATCTGAGAGCCCTCCAGCTAGCTCC 565
Db      506 GAGGAGTGAAGGTGAATACCAGTCCCTGCAATCTGAGAGCCCTCCAGCTAGCTCC 565
QY  1726 GAAAAGAGCTTCGCGGAGAGCAGTGTAGAGCTTTCAACAGGCTCAACACAGCACCAC 1785
Db      1726 GAAAAGAGCTTCGCGGAGAGCAGTGTAGAGCTTTCAACAGGCTCAACACAGCACCAC 1785
QY  566 GGCAAGAGCTTCGCGGAGAGCAGTGTAGAGCTTTCAATGCTTCAACAGCAGACAGCAG 625
Db      566 GGCAAGAGCTTCGCGGAGAGCAGTGTAGAGCTTTCAATGCTTCAACAGCAGACAGCAG 625
QY  1786 CGGCTCACTTCGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 1845
Db      1786 CGGCTCACTTCGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 1845
QY  626 CGGCTCACTTCGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 685
Db      626 CGGCTCACTTCGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 685
QY  1846 TGCAAGCTCATCTGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 1904
Db      1846 TGCAAGCTCATCTGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 1904
QY  686 TGTAGCTCATCTGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 744
Db      686 TGTAGCTCATCTGCGGAGAGCAGTGTAGAGCTTCGCGGAGAGCAGTTCGCGGAGAGCAG 744
QY  1905 GGTGGA-CGGCAGCCTGTG-CTCTGCTGAGTCCAGCTCCGCTGTGTCGAAGGCAAG-TG 1961
Db      1905 GGTGGA-CGGCAGCCTGTG-CTCTGCTGAGTCCAGCTCCGCTGTGTCGAAGGCAAG-TG 1961
QY  745 GGTGAGCCGAGTACGCTGTGTACACCTGACCTCCAGCTCCGCTGTGTGTGTCGAAGGCAAGT 804
Db      745 GGTGAGCCGAGTACGCTGTGTACACCTGACCTCCAGCTCCGCTGTGTGTGTCGAAGGCAAGT 804
QY  1962 CATCAA-GGCTGCTGTGATGAGGAGACCTGGGCTCAAGAA 2000
Db      1962 CATCAA-GGCTGCTGTGATGAGGAGACCTGGGCTCAAGAA 2000
QY  805 CATCACAGGGGCTGTGCGAAGCGCAAACTGGGCTCAAGCA 845
Db      805 CATCACAGGGGCTGTGCGAAGCGCAAACTGGGCTCAAGCA 845

RESULT 3
LOCUS   BG921734
DEFINITION BG921734 681 bp mRNA linear EST 05-JUN-2001
          60282556F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:454456 5',
          mRNA sequence.
ACCESSION BG921734
VERSION   BG921734.1 GI:14302210
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS   1 (bases 1 to 681)
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: Jeffrey E. Green, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L14M0915 row: 1 column: 17
           High quality sequence stop: 680.
           Location/Qualifiers
             1..681
               /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone_image="454456"
               /lab_host="NCI_CGAP_Mam6"
               /sex="female, virgin"
               /tissue_type="infiltrating ductal carcinoma"
               /dev_stage="5 months"
               /lab_host="DH10B"
               /note="Organ: mammary; Vector: pCMV-Sport6; Site:1: SalI;
               Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT."

```


Query Match	13.7%	Score 391.4	DB 10	Length 522
Best Local Similarity	86.7%	Pred. No. 6.4e-74		
Matches 431	Conservative	0	Mismatches 66	Indels 0
				Gaps 0

EST discovery in swine
Unpublished (2000)

COMMENT Contact: Smith RPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail@marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACAT
BACKWARD: GTTTCACGACGACG
Plate: 43 row: P column: 17
Seq primer: ATTAGTGCATATGAC.
Location/Qualifiers
1. 544
/organism="Sus scrofa"
/db.xref="taxon:9823"
/clone.lib="MARC 2Pig"
/tissue.type="pooled"
/lab.host="DH10B"
/note="Vector: PCMV SPORT6; Site.1: NotI; Site.2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 117 a 154 c 169 g 104 t

ORIGIN

Query Match 12.3%; Score 349.6; D3 12; Length 544;
Best Local Similarity 88.6%; Pred. No. 6.4e-55;
Matches 379; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 559 GCCCTGACCTTACAAAGCGCGGCGGCTTCGGGAGAGCTAGCGCGCGAGG 618
Db 544 GCCCTGACCTTACAAAGCGCTCCGGATAGGCTTAGGAGAGCTGCAAGCGCGAGG 485

QY 619 TCTGGCGCGCAAGCGTTTCTATCCGCGGCTACGTGAGAGCGTGTGTGCG 678
Db 484 TCCGGCGCGCAAGCGCTTCGTCTATCCGCGATAGTGAACGCTGTGTGTGCA 425

QY 679 GAGAGTCATGTGTCAGTTCCAGCGCGGAGCCTTGACATTTATCTGTGACGCTGCTG 738
Db 424 GAGAGTCATGTGTCAGTTCCAGCGCGGAGCCTTGACATTTATCTGTGACGCTGCTG 365

QY 739 GCAACGGCGCGGAGCTTACCGCCATCCGACATCTTCAACCCATCAACATCGTTGTG 798
Db 364 GCAACGGCGCGGAGCTTACCGCCATCCGACATCTTCAACCCATCAACATCGTTGTG 305

QY 799 GTCAAGTGTGCTCTTAAAGATCGTGAATCCGGCGCCCAAGTACCGGCAATCGCGCC 858
Db 304 GTCAAGTGTGCTCTTAAAGATCGTGAATCCGGCGCCCAAGTACCGGCAATCGCGCC 245

QY 859 CTGACGCTGCGCAACTTCTGTGCTTGCGGCAAGAAAGCTGACAAATGATGACAAAGCAG 918
Db 244 TTGACGCTGCGCAACTTCTGTGCTTGCGGCAAGAAAGCTGACAAATGATGACAAAGCAG 185

QY 919 CCGAGTACTGGGACACTGACATCTCTTACACAGGACGACCTGTGGAGCCGCAAGC 978
Db 184 CCGAGTACTGGGACACTGACATCTCTTACACAGGACGACCTGTGTGTGACCTCT 125

QY 979 TGTGACAC 986
Db 124 TTGAGACC 117

RESULT 7
LOCUS BE281680 627 bp mRNA EST 26-OCT-2000
DEFINITION 601099657f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491991 5',
ACCESSION BE281680
VERSION BE281680.1 GI:9156727
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 627)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM8537 row: a column: 16
High quality sequence stop: 530.
Location/Qualifiers
1. 627
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone.lib="IMAGE:3491991"
/clone.lib="NCI_CGAP_Lu29"
/tissue.type="spontaneous tumor, metastatic to mammary."
/lab.host="DH10B"
/note="Organ: Lung; Vector: PCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 153 a 159 c 177 g 138 t

ORIGIN

Query Match 11.9%; Score 339.4; DB 10; Length 627;
Best Local Similarity 83.6%; Pred. No. 1e-62;
Matches 504; Conservative 0; Mismatches 86; Indels 13; Gaps 10;

QY 1850 AGCTCATTCGCGAGGACATGACCTGCTATGTCGTCGACACCAAGGAGTGGT 1908
Db 1 AGCTCATTCGCGAGGACATGACCTGCTATGTCGTCGACACCAAGGAGTGGTGG 60

QY 1909 GACGCGAGCGTGTCTCTCTGAC--TCCACTTCGCTCTGTGTCAGGACAGTGCATCAA 1967
Db 61 GACGCGAGCGTGTCTCTCTGAC--TCCACTTCGCTCTGTGTCAGGACAGTGCATCAA 120

QY 1968 -GGCTGCTGTGTGAGAACCTGG--GCTCCAGAGAGATTCGACAA--GTGTGGGCTGTG 2024
Db 121 TGCGTGGCTGCGAGGAGATCTGTGCTCCAGAGAGATTCGACAA--GTGTGGGCTGTG 180

QY 2025 TGGGGAGACAAATA--GAGCTCAGAGAGAGTGTGACTCTTCACCAAGCCATGCATG 2083
Db 181 TGGTGGAGACAAATA--GAGCTCAGAGAGAGTGTGACTCTTCACCAAGCCATGCATG 240

QY 2084 GCTACAAATTCGAGGAGGACATCCCGAGGAGGCTCAAGATGACATCCCGAGGAGG 2143
Db 241 GCTACAAATTCGAGGAGGACATCCCGAGGAGGCTCAAGATGACATCCCGAGGAGG 300

QY 2144 GTTACAAAGGCTGATGGGGATGACAACTACCTGCTGTGAAGAACAGCCAGCAAGT 2203
Db 301 GTTACAAAGGCTGATGGGGATGACAACTACCTGCTGTGAAGAACAGCCAGCAAGT 360

QY 2204 ACCTGCTCAAGGAGGATTCGAGTGTGCGGCGGTGAGAGGAGGAGCTGTGTGTA--GG 2261
Db 361 ACCTGCTCAAGGAGGATTCGAGTGTGCGGCGGTGAGAGGAGGAGCTGTGTGTA--GG 420

QY 2262 CACTCTGCTGCGTACAGCGCGAGGACACAGCGGTGAGAGGCTTCGACGCTTCGCGGCC 2321
Db 421 CACTCTGCTGCGTACAGCGCGAGGACACAGCGGTGAGAGGCTTCGACGCTTCG-C GAC 479

QY 2322 CATCTGAGAGCGGCTGACCGGTGAGGCTCTCCGTGGGGAAGATGACACCGCCCGGCT 2381
|||||


```

/clone="IMAGE:5176476"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pcmv-sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age 69. Library is
oligo-dt primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC library."

```

BASE COUNT 169 a 204 c 225 g 140 t

ORIGIN

Query Match 9.2%; Score 263.4; DB 13; Length 738;

Best Local Similarity 63.1%; Pred. No. 2.4e-46;

Matches 440; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

```

QY 1715 GCTCAGCTCCGGAAGAGTTCCGGAGGACAGTGTAGGCTTTCACAGGCTCAACC 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 GCGCCCTGAGGGAAGAGCTTCAGGACAGAGTGTAGAAATATATAGCTTCAATT 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1775 ACAGCACCACCGCTCCTCTCCGCGATGGATGGTCCCAAGTACTCCGCGTCTC 1834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 ACATGACATGACGGAATCT---CCTGAGTGGGTCCCAAGTATGCTGGGTTCC 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1835 CCGGGAACAGTGCAGCTCATCTGCGGAGCCATGGCTACTTCTATGCTGG 1894
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 CCGGGAACAGTGCAGCTCATCTGCGGAGCCATGGCTACTTCTATGCTGG 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1895 CACCAAGTGTGAGCAGCAGCTGCTCTCTGCTACCTCCACCTGCTGTCAG 1954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 AGGCAAGTGTGATGATGACACCTGTGTGGCAGCAAAACACTGGCACTGTGCTG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1955 GCAGTGCATCAAGGCTGCTGTGATGGACACTGGGCTCCAAAGAGATTCGACAGT 2014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 GCGAGTGTGAAAGCGCGGTGACCATGTGTGATGCTGCTCGAAGCTGGACAAAT 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2015 GTGGGTGTGTGGGAGACAAATAGAGTGCAGAAAGTGTGAGTCTTCACCAAGC 2074
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 GCGGCTGTGTGGGAGAAAGCACTCCTGCAAGAAAGTGTGCGGTTCCCTCACCCCA 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2075 CCATGATGCTACATTTGTGTGGCCATCCCGGAGGCGCTCAAGCATGCATCC 2134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 CCATTTATGCTACATGATGATGATGATGATGATGATGATGATGATGATGATG 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2135 GCCACGCGGTGTCAAAAGGCTGATGGGATGACAACTACTGCTCTGAGAACAGCC 2194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 AGCAGCGGAGCCACCGGGGTGTGCAAGATGGAGATGAGTGGCGCTGAGAGAGCTG 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2195 AAGGCAAGTACCTGCTCAACGGGCAATTTGTGTGGCGGTGAGCGGACCTGGTGG 2254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 ATGGGCTGTGCTCAAGGAGGCAACCTGCTCTGCTCAAGGAGGCACTTGG 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2255 TGAAGGAGCACTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 TGAAGGAGCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2315 CCGGAGCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2372 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10
LOCUS B1733795 746 bp mRNA linear EST 20-SEP-2001
DEFINITION 603351940F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5359712 5',

```

ACCESSION mRNA sequence.
VERSION B1733795
KEYWORDS B1733795.1 GI:15710808
SOURCE EST.
ORGANISM house mouse.
MUS musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 746)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.jnl.gov
Plate: LHM11914 row: 0 column: 09
High quality sequence stop: 746.
Location/Qualifiers
1. 746
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5359712"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/notes="Organ: eye; Vector: pcmv-sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

```

BASE COUNT 213 a 160 c 216 g 156 t 1 others

ORIGIN

Query Match 8.9%; Score 254.8; DB 13; Length 746;

Best Local Similarity 61.3%; Pred. No. 1.7e-44;

Matches 447; Conservative 0; Mismatches 273; Indels 9; Gaps 2;

```

QY 1463 CCATGGACACAGCTGTGGCGAGGCAAGCTGCTCAAGGGGCTGCGTGGAGAC 1522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CAGATGGACACAGCTGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1523 ACAACCTCAACAA-----GCACAGGAGTGGATGTTCTGCGGCAAAATGGGATCTATG 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CAGACATGAAGCAATTTGTACTCTGTCATGAGAAAGCTGGGAGACATGGGACCGTGG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1577 GCGGCTGTGCGGACATATGTTGGGGGCGGTGACACTGGCCAGAGGAGGACACCAAGC 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GAGACTGCTCAAGAACCTGTGTGTGGGTGAGTCAATACATGAGAGATGTGACAAAC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1637 CCACCCCTGCAACGGGGGCAAGTACTGCGAGGAGTGGAGGTGAATACCATCTGCA 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CACTGCCAAGAAAGAGGAGGAGTACTGTGAAGCAAGATCGCTACAGGCTCTGTA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1697 ATGTGAGCGCTGCCACACTCAGCTCCGGAAGAAAGCTTCCGAGAGAGAGAGTGAAG 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ACATCGAGGAGCTGTCCAGAC---AATTAAGGAAAGCTTCAAGAGGAGGAGTGAAG 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1757 CTTTCAAGGAGTACAACACACAGCACCACCGGCTCACTGTGCGGCTGGCATGGTGGCCA 1816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 CCGACATAGATTTTCCAAAGCTCTGTTGGAAGATGAGCCACATGTATAGAGGAGACCA 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1817 AGTATCGGCGGTGTCCCGGAGCAAGTGAAGTATATGCGGAGGACCATGGACTG 1876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 AGTACCGCGCGCTGTCCGCAAGGAGCAGGTGCAACCTCACTGTGAAGCCAAAGCAATTG 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1877 GCATCTTATGATGCTGGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	418	GCTACTTTTTCGCTTTCACGCCCAAGGTTTGAATGGCACTCCCTGTATGTCAGACACTCA	477
QY	1937	CCTCCGTCGTGTGCCAAGGCAAGTGCATCAAGCGCTGCTGTGATGGGAACTCTGGGCTTCCA	1996
Db	478	CCCTGTCTGTGTGCCAAGGGCAGTGTGTGAAGACGTGGCTGTGATGGCACTCATATGACTCA	537
QY	1997	AGAAAGATTTGCAACAAGTGTGGGTGTGTGGGGGAGACATTAAGAGCTGCAGAAAGTGA	2056
Db	538	AAAAGAASTTGGATTAAGTGTGGCGTTTGTGGAGGAAACGTTCCACATCAAGAAAGATGT	597
QY	2057	CTGACCTTTTACCAAGGCCATGCATGGCTGTCAATTTCGTGTGGCCATTCGCCGAGGCG	2116
Db	598	CAGGAATGTCACTAGTACAAAGACCTGGGTATCATGACATGTGTCAATTCTCTGTGGAG	657
QY	2117	CCTCAAGCATGCATCCGCGAGCGCGGTTACAAAGGGCTATCGGGGATGCAACACTAC	2176
Db	658	CCACCACATGTAAAGTGAACATCGGATTCAAAGGGGCTCCAGAAACAATGGACACTTTC	717
QY	2177	TGGCTGTGA	2185
Db	718	TGGCTATTTA	726

RESULT	11
LOCUS	B8220373
DEFINITION	B8220373 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA A530059019 3' similar to AF140673 Mus_musculus putative secreted metalloprotease ADAMT5 (Adamts5) mRNA, mRNA sequence.
ACCESSION	B8220373
VERSION	B8220373.2 GI:16353495
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 651)
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., et al.

TITLE
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 JOURNAL
 Unpublished (2001)
 COMMENT
 On Jun 30, 2000 this sequence version replaced gi:8885326.

On Jun 30, 2000 this sequence version replaced g1:8885326.
 Contract: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitoh-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel.: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp,
 URL: <http://genome-gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Wataniki, M., Yoneeda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga
 S., Kawai, I., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
 Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a

BASE COUNT	ORIGIN	FLC 1
160 a	167 c	180 g
154 t		
<p>Query Match 8.6%; Score 245.8; DB 10; Length 661; Best Local Similarity 60.9%; Pred. No. 1,4e-42; Matches 400; Conservative 0; Mismatches 257; Indels 0; Gaps 0;</p>		
QY	861 GACGGTGGGCAACTCTGTGGCTGGCCAGAAAGAGCTGAACAAGTAGAGTGAACAGCACC	920
DB	4 GACCTTCAAGAACTTTTGGCAATGSCACGACCAACATATAAGCAGTAGAGGATGATCAGCA	63
QY	921 CGAGTACTGGGACACTGCGATCTTCTTCAACGACGAGACCTGTGTGAGCCACCACTG	980
DB	64 AGAGCACTAGCTGACGACCATCTTTCACCCGAGAGGATTTATGTGGCATCATTTCAAG	123
QY	981 TGACACCTGGGCAATGGCTGATGTGGGTACATATGTGTGATCCCCAGAGAGATGGCTGT	1040
DB	124 TGACACCTGGGAAATGGACGACGTTTGGACCACTATATTTCTCCGGAGCGACGCTGTGCA	183
QY	1041 CATTTAGAGCAATGGGCTTCCATTCAGACCTTTCACCACTGCGCCACGAGACTGGCCAGCTGT	1100
DB	184 GATTAGATGATGTGGCTTCATGCAAGCCTTTCACCTGTGGCTCATGAAATTTGGCATCTTACT	243
QY	1101 CAACATGCCCATGACAAATGTGAAGTGTGTGAGGAGCTTTGGGAACCTCCGAGCCAA	1160
DB	244 TGGCTTTCTCAATGACGATTCGAAATTCGTGAAGAGAACTTCGTTACTACAGAGACAA	303
QY	1161 CCACATGATATCTCCCGACCTTCATCCAGATGACGCTGCAACCCCTGGTCA6CTTGAG	1220
DB	304 GGGTTTAATGTCTTCAATCTCTTACGAGCATGATGATCAGACCTGGTCCAAATATGAC	363
QY	1221 TGTGTCATCATTCACGACTCTCTGTGACAGGGGGAGCGGTGATGCTCTGTGGACCAAC	1280

Db	364	GTACGACCCATCATCAGAAATTCTCGATGATGTATCATGTAAATTGTTGCTAGACTAC	423
QY	1281	CAGAACCCATCTCCCTCGCCGAGAGATCGCCGGGCGGACCTACACCCATGACGACCA	1340
Db	424	ACGAGACAGATTTTGGGCTCCCGAGGAACTCCAGGACACTCGATGATGCCACCCACA	483
QY	1341	GTGCGAGCTGCTTTTGGCGTGGGCTCCAGGCCCTGTCTTACATGACAGTACTGCACAA	1400
Db	484	GTGCACCTTACATTTTGGGCTGAGTACGTGCTGCTCCCTGCATGAGATGCTGTGGCG	543
QY	1401	GCTGTGGGACCGGGGAAGGCCAAGGAGACGATGTGTCCGACACCGGCACTTCCCTG	1460
Db	544	GCTGTGGTGTGTGTGGTGGGCCAAGGCCAAATGGTGTCTCTGACCAGAAAGCTGCCG	603
QY	1461	GGCCGATGGCAACAGCTGTGGCCAGGGCAAGCTCTGTCTCTCAANAAGGGCGCTCGTGG	1517
Db	604	TGTGGACGACCTCCCTGTGTGGGAAGAAAGATGTCTCTTCACAGGCAAAATGTGTGA	660

RESULT 12	
BB651049	
LOCUS	BB651049
DEFINITION	BB651049 RKEN full-length enriched, 0 day neonate cerebellum Mus musculus cDNA clone C23003J22 5' mRNA sequence.
ACCESSION	BB651049
VERSION	BB651049.1 GI:16485303
EST 26-OCT-2001	

ORGANISM

.....

AUTHORS

TITLE
JOURNAL,
COMMENT

Eumkariota; Euteleostei: Chordata; Craniata; Vertebrata; Euteleostomi; *Mammalia*; *Eutheria*; *Rodentia*; *Sciurognathi*; *Muridae*; *Murinae*; *Mus*.
 1 (bases 1 to 660)
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanaigaki,T., Hara,A.,
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasak,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,K., Takahashi,F., Takeda,Y., Tanaka,T., Toyu,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshinhide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome.res@gsc.riken.go.jp,
 url: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
 M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 wagi,K., Fujisake,S., Inoue,K., Togawa,Y., Ikawa,M., Ohara,E.,
 Matshiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,T., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamataka,I., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)

```

FEATURES
source      1. .660
location/Qualifiers
e mouse     tissues.
further details.
Please visit our web site (http://genome.gsc.riken.go.jp) for

```

BASE COUNT ORIGIN	160 a	166 c	185 g	149 t
----------------------	-------	-------	-------	-------

[illegible]


```

RESULT 14
B0930852          951 bp      mRNA      linear      EST 20-AUG-2002
LOCUS             AGNCOURT.8920927 NCI CGAP Co24 Mus musculus cDNA clone
DEFINITION        IMAGE:6395437 5', mRNA sequence.
ACCESSION         B0930852
VERSION           B0930852.1  GI:22345863
KEYWORDS          EST.
SOURCE            house mouse.
ORGANISM          Mus musculus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE         NIH-MGC http://mgc.nci.nih.gov/
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: c9apds-remail.nih.gov
                  Tissue Procurement: The Cepko Laboratory
                  cDNA Library Preparation: Life Technologies, Inc.
                  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM3890 row: b column: 14
                  High quality sequence stop: 644.
                  Location/Qualifiers
                    1..951
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone_image="6395437"
                     /clone_lib="NCI CGAP Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI CGAP Library."
BASE COUNT        189 a      143 c      180 g      135 t      304 others
ORIGIN
Query Match      8.4%; Score 240.2; DB 14; Length 951;
Best Local Similarity 62.1%; Pred. No. 2.6e-41;
Matches 396; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

OY 1549 GGTTCCTGGGCGCAATGGATCCTATGGCCCTGCTCGGCGACATGTGTCGGGGCGCTG 1608
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GGAAGCTGGGGACCATGGGGACCTGGGGAGACTGCTCAGAACCTGTGCTGGTGGAGT 60

OY 1609 CAGTGGCGGAGGAGGAGTGGACACCCGACCCCTGCCAAGGGGGCAAGTATCTCGAG 1668
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CATTACATATGAGATGTGACACACCTCCCAAGAGAGGAGGAGTACTGTGA 120

OY 1669 GAGTGGAGGTGAATACCATCTCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGA 1728
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GGCAACAGATGCCCTACAGTCTCTAATCATGTGAGAGACTGTCAGAC--ATAACGGA 177

OY 1729 AAGAGCTTCGGGAGGAGCACTGTGAGGCTTTCAAGCGCTACACCAAGCAACCGG 1788
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 AAAAGCTTCAGAGAGGAGCACTGTGAGGCGCACAAATGATTTTCCAAGCTTCTTTGGG 237

OY 1789 CTCACCTCGCGCGTGGATGGGTGCCCAATGACTCGGCTGTCTCCCGGCAAGTGC 1848
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 238 AATGAGCCCACTGTAGATGTGACACCCAAATAGCCGCGCTCTGCCAAAGGAGAGTGC 297

OY 1849 AAGCTCATCTGCGCAATGAGCACTGCTTCTATGTGCTGCGACCCCAAGTGTG 1908
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 AAGCTCATCTGTGAAGCCAAAGGCACTTCTTCTTACAGCCCAAGTGTGA 357

OY 1909 GAGGCAAGCTGTCTCTCGATCGACTCGGTGTGTGTGTCAAGCAAGTGCATCAG 1968
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 GATGCACTCCCTGTAGTCCAGACTCTGTCTGTGTGTGTGAGGAGGAGTGTGTGANA 417

```

```

OY 1969 GCTGCTGTATGGGAACCTGGGCTCCAGAGAGATTCAGACAGTGTGGGTGATGGG 2028
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 418 GCTGCTGTATGATGCATCATAGACTCCAAAGAAATTTTGATTAAGTGTGGCTTGTGA 477

OY 2029 GGAACATATAGACCTGCAAGAAAGGTGACTGTGACTTTTACCAAGCCCATGCTGATC 2088
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 478 GGAACAGTTCCACATGCAAGAAAGATGTCAGATGATCAAGACCTGAT 537

OY 2089 AATTCTGTGTCATCCCGGAGCGCTCAACATGACATCGGCGCGCTTAC 2148
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 538 CATACATTTTCACATCTCTGCTGAGGACACCAACATTTGATGATGCAATCA 597

OY 2149 AAGGCTGATCGGAGTGCACACTACCTGCTCTAA 2166
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 598 AGGGGTCCAGAAACATGCGACCTTCTGTATTAA 635

RESULT 15
BFO33017          945 bp      mRNA      linear      EST 20-OCT-2000
LOCUS             601455701F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3859104 5',
DEFINITION        mRNA sequence.
ACCESSION         BFO33017
VERSION           BFO33017.1  GI:10740729
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE         NIH-MGC http://mgc.nci.nih.gov/
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: c9apds-remail.nih.gov
                  Tissue Procurement: DCTD/DRP
                  cDNA Library Preparation: Life Technologies, Inc.
                  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM592 row: b column: 01
                  High quality sequence stop: 642.
                  Location/Qualifiers
                    1..945
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_image="IMAGE:3859104"
                     /clone_lib="NIH_MGC_66"
                     /tissue_type="adenoecarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pCMV-Sport6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.8 kb. Library constructed by Life
                     Technologies."
BASE COUNT        220 a      246 c      326 g      153 t
ORIGIN
Query Match      8.1%; Score 229.8; DB 12; Length 945;
Best Local Similarity 61.6%; Pred. No. 4.4e-39;
Matches 421; Conservative 0; Mismatches 252; Indels 10; Gaps 3;

OY 1534 AAGACAGAGGTGATGTTCTTGGGCAATGGGATCCCTATGCGCCCTGCTGGGCA 1593
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16 AGGCCCTGTGTAATGAGGCTGGGACCTGTGGGAGAAATGTTCTGAGGAC 75

OY 1594 TGTGTGGGGGCGTGCAGTGGCCAGAGGAGTGCACCAACCCACCTGCAACGG 1653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 TGTGAGAGAGGATGACGTTTTCACACCGTGAATGCAAGACCCGAGCTCAGAAATGA 135

OY 1654 GGCAGTACTGCGAGGAGTGAAGGTAATACGATCTCGAATCTGAGAGCCCTGCCCC 1713

```

```
Db 136 GGAAGATACGCTGGGTGGAGAGCCAGTACAGTCAAGCAGAGATGCCCC 195
QY 1714 AGCTCAGCCTCCGGAAAGACCTTCCGGAGAGACAGTGTAGAGCTTCAACGGCTACAAAC 1773
Db 196 -----CCTGACGCGGAAAGCTTCAGGAGACAGAGTGTAGAGATATATGCTACAAAT 249
QY 1774 CACAGACCAACCGGCTCACTCCGCCGTGGGATGGGTGCCAAGTACTCCGGCGTGTCT 1833
Db 250 TACACTGACATGGACGGGAATCT--CCTGCACTGGGTCCCAAGTATGCTGGGGTGTCC 306
QY 1834 CCCCAGGACAAAGTCAAGCTCATCTGCTCCGACCAATGCACTGACTACTTCTATGTGTG 1893
Db 307 CCCCAGGACCGCTGCAAGTTGTTCTGCCGAGCCCGGGAGAGCGAGTTCAAAGTGTTC 366
QY 1894 GCACCCCAAGTGTGTGACGCGACAGCTGTGCTCTCTGACTCCACCTCCGTCGTGTCCAA 1953
Db 367 GAGGCCAAGTGTATGTGACACCTGTGTGGGCCAGAAACACTGGCCATCTGTGTCCGT 426
QY 1954 GGCAGTGCATCAAGGCTGGCTGTGTATGGGAACCTGGGCTCCAAAGAGATTCGACAAAG 2013
Db 427 GGCAGTGTGTCAAGGCCGGCTGTGACCATGTGTGACTCGCTCGGAAGCTGGACAAA 486
QY 2014 TGTGGGTGTGTGGGGAGACAAATTAAGAGCTGCAGAGAGTGTGACTGACTCTTACCCAAG 2073
Db 487 TCGGGGTGTGTGGGGCAAAAGGCAACTCCTGCAGAGAGGTCTCGGGTCCCTCACCCCC 546
QY 2074 CCGATGATGGCTACAAATTTCTGTGTGGCCATCCCGCAGGCCCTCAACCATCGACATC 2133
Db 547 ACCAATTATGTGCTACATGACATTTGCACCATCCAGCTGTGTGCCACTAATATTTAGCGTG 606
QY 2134 CGCCAGCGCGTTACAAAGGGGTGATCGGGGATGATGACACTACTGTGCTCTGAAGAAGAGC 2193
Db 607 AAGCAGCG -GAGCACCCGGGTGTGCAAGAGATGGGAAGTACTACTGGGCGTTGAAGAGGT 665
QY 2194 CAAGGCAAGTACTGCTCAACGG 2216
Db 666 GATGGGCACTACTGCTCAAGGG 688
```

Search completed: May 15, 2003, 22:52:54
Job time : 3616 secs


```

QY 841 GTACCGGCAATGGGGCCCTGACGCTGCGCACTTCTGTGCTGGCGAAGAAAGCTGAAAC 900
DB 841 GTACCGGCAATGGGGCCCTGACGCTGCGCACTTCTGTGCTGGCGAAGAAAGCTGAAAC 900
QY 901 AAAGAGTGCACAAAGCAACCCGAGTACTGAGCACTGCTCCATCCCTCTTACACAGGCAAGAC 960
DB 901 AAAGAGTGCACAAAGCAACCCGAGTACTGAGCACTGCTCCATCCCTCTTACACAGGCAAGAC 960
QY 961 CTGTGTGAGGACCAACCTGTGACACCCCTGGGCACTGCTGATGTGGCTACATGTGTGAC 1020
DB 961 CTGTGTGAGGACCAACCTGTGACACCCCTGGGCACTGCTGATGTGGCTACATGTGTGAC 1020
QY 1021 CCCAAGAGAGAGCTGCTGTGATGAGAGAGATGGGCTTCATCAGGCTTCACACAGCTGAC 1080
DB 1021 CCCAAGAGAGAGCTGCTGTGATGAGAGAGATGGGCTTCATCAGGCTTCACACAGCTGAC 1080
QY 1081 CACGAGCTGGGCAACGCTTCAACATGCCCCATGACAAATGTGAAGGCTGTGAGAGAGT 1140
DB 1081 CACGAGCTGGGCAACGCTTCAACATGCCCCATGACAAATGTGAAGGCTGTGAGAGAGT 1140
QY 1141 TTTGGAGAGCTCCGAGGCAACATGATGTCGGGACCCCTCATCCAGATCGACCGTGGC 1200
DB 1141 TTTGGAGAGCTCCGAGGCAACATGATGTCGGGACCCCTCATCCAGATCGACCGTGGC 1200
QY 1201 AACCCCTGGTACGCTGAGTGTGCTGCAATCACCAGACTTCTGTGACAGCTGGGCAAGGT 1260
DB 1201 AACCCCTGGTACGCTGAGTGTGCTGCAATCACCAGACTTCTGTGACAGCTGGGCAAGGT 1260
QY 1261 GACTGCTCTGTGAGCAACCCAGCAAGGCCATCTCCCTGCGGAGATCTGCGGCGGCGC 1320
DB 1261 GACTGCTCTGTGAGCAACCCAGCAAGGCCATCTCCCTGCGGAGATCTGCGGCGGCGC 1320
QY 1321 AGCTACACCTGAGGCAAGTGTGAGAGCTGCTTTGGCGTGGCTTCAAGCCCTGTGCT 1380
DB 1321 AGCTACACCTGAGGCAAGTGTGAGAGCTGCTTTGGCGTGGCTTCAAGCCCTGTGCT 1380
QY 1381 TGCATGCACTGACCAACCACTGTGTGACACCGGGAAGCCAAAGGCAAGATGTGTG 1440
DB 1381 TGCATGCACTGACCAACCACTGTGTGACACCGGGAAGCCAAAGGCAAGATGTGTG 1440
QY 1441 CAGACCCGCACTTCCCTGGGCGAGTGGACAGCTGTGGCGAGGCGCAAGCTGTGCT 1500
DB 1441 CAGACCCGCACTTCCCTGGGCGAGTGGACAGCTGTGGCGAGGCGCAAGCTGTGCT 1500
QY 1501 AAAAGGGCTGTGTGAGAGACACAACCTCAACAAGCAAGGCTGTGCTGTGCTG 1560
DB 1501 AAAAGGGCTGTGTGAGAGACACAACCTCAACAAGCAAGGCTGTGCTGTGCTG 1560
QY 1561 AAATGGATCCCTATGAGCCCTGTGCTGCGGCACTGTGTGGGCGCTGACAGCTGTG 1620
DB 1561 AAATGGATCCCTATGAGCCCTGTGCTGCGGCACTGTGTGGGCGCTGACAGCTGTG 1620
QY 1621 AGGCACTGACCAACCCCAACCCCTGCAAGGCGGCAAGTACTGCGAGGAGTGAAGGTG 1680
DB 1621 AGGCACTGACCAACCCCAACCCCTGCAAGGCGGCAAGTACTGCGAGGAGTGAAGGTG 1680
QY 1681 AAATACCATCTGCAATCTGAGAGCCCTGCGCCAGCTGAGCTGCGGAAAGAGCTTCG 1740
DB 1681 AAATACCATCTGCAATCTGAGAGCCCTGCGCCAGCTGAGCTGCGGAAAGAGCTTCG 1740
QY 1741 GAGGAGCAAGTGTGAGGCTTTCAAGGCTACCAACCAAGCAACACCGGCTCACTTCG 1800
DB 1741 GAGGAGCAAGTGTGAGGCTTTCAAGGCTACCAACCAAGCAACACCGGCTCACTTCG 1800
QY 1801 GTGGATGGGCGCAAGTACTGCGGCGTGTCCCGGGAAGCAAGTCAAGCTCATTCG 1860
DB 1801 GTGGATGGGCGCAAGTACTGCGGCGTGTCCCGGGAAGCAAGTCAAGCTCATTCG 1860
QY 1861 CGAGGCAATGGCACTGCTACTTCTATGTGCTGACCAAGGAGTGTGAGAGGCAAGCTG 1920
DB 1861 CGAGGCAATGGCACTGCTACTTCTATGTGCTGACCAAGGAGTGTGAGAGGCAAGCTG 1920
QY 1921 TGCCTCTGACTGCACTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
DB 1921 TGCCTCTGACTGCACTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980

```

```

DB 1921 TGCCTCTGACTGCACTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGGAAACCTGGGCTCCAAAGAGAGATTTGACAAAGTGTGTGGGTGTGTGGGAGACAAATAG 2040
DB 1981 GGGAAACCTGGGCTCCAAAGAGAGATTTGACAAAGTGTGTGGGTGTGTGGGAGACAAATAG 2040
QY 2041 AGCTGCAAGAGTGTGACTGTGACTTCTTACCAAGCCCATGATGGCTACAAATTTGCTGTG 2100
DB 2041 AGCTGCAAGAGTGTGACTGTGACTTCTTACCAAGCCCATGATGGCTACAAATTTGCTGTG 2100
QY 2101 GCGATCCCGGAGGCGCTCAAGCATGACATCCGCGAGCGCGGTATCAAAAGGCTGATC 2160
DB 2101 GCGATCCCGGAGGCGCTCAAGCATGACATCCGCGAGCGCGGTATCAAAAGGCTGATC 2160
QY 2161 GGGATGACAACTGCTGTGCTGTGAGAGACGCCAAGGCAAGTGTGCTCAACGGGCTAT 2220
DB 2161 GGGATGACAACTGCTGTGCTGTGAGAGACGCCAAGGCAAGTGTGCTCAACGGGCTAT 2220
QY 2221 TTGCTGTGCTGGGCTGTGAGAGCGGACCTGTGTGAAGGCGAGTGTGCTGTGCTGACAGC 2280
DB 2221 TTGCTGTGCTGGGCTGTGAGAGCGGACCTGTGTGAAGGCGAGTGTGCTGTGCTGACAGC 2280
QY 2281 GGCAGGCGGACAGCGGTGTGAGAGCGCTGACAGCTTCCCGGCGCCATCTGAGCGCTGAC 2340
DB 2281 GGCAGGCGGACAGCGGTGTGAGAGCGCTGACAGCTTCCCGGCGCCATCTGAGCGCTGAC 2340
QY 2341 GTGGAGGCTCTCTCGGTGGGGAAGATGACACCGCGCGGCTCGGTACTCTTCTATCTG 2400
DB 2341 GTGGAGGCTCTCTCGGTGGGGAAGATGACACCGCGCGGCTCGGTACTCTTCTATCTG 2400
QY 2401 CCCAAAGAGCTCGGAGAGACAAAGTCTCTATCCCAAGAGCCCGGAGACCTCTGTCT 2460
DB 2401 CCCAAAGAGCTCGGAGAGACAAAGTCTCTATCCCAAGAGCCCGGAGACCTCTGTCT 2460
QY 2461 TTGCAACAACAGCTGTCTGAGCTGTCTTCAACAGTGTGAGAGCGGAGCAACAGGCGCT 2520
DB 2461 TTGCAACAACAGCTGTCTGAGCTGTCTTCAACAGTGTGAGAGCGGAGCAACAGGCGCT 2520
QY 2521 GCACGCTGGTGTGCTGAGCTGTGGGCGCTGTGCTGCGGAGGCGAGCTGTGCTGTGATGA 2580
DB 2521 GCACGCTGGTGTGCTGAGCTGTGGGCGCTGTGCTGCGGAGGCGAGCTGTGCTGTGATGA 2580
QY 2581 AAGCGGCGGTGTGAGCTGTGGGCGCTGTGCGGAGGCGAGGCTGTGCTGTGATGA 2640
DB 2581 AAGCGGCGGTGTGAGCTGTGGGCGCTGTGCGGAGGCGAGGCTGTGCTGTGATGA 2640
QY 2641 GCCATGCGGCGGTGTGAGACACAAGCTGTGCGGAGGCGCTGTGCGGAGGCTGTGATGA 2700
DB 2641 GCCATGCGGCGGTGTGAGACACAAGCTGTGCGGAGGCGCTGTGCGGAGGCTGTGATGA 2700
QY 2701 GCGTGTGACCTGTCTCAAGAGCTGTGGGCGGCGGAGTTCAGAGGCGCTCACTCAAGTGT 2760
DB 2701 GCGTGTGACCTGTCTCAAGAGCTGTGGGCGGCGGAGTTCAGAGGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCAAGGAGCGGCTGTGCGGCGGAGGAGCTGTGCACTTGTGACCGCAAGCCCGCAG 2820
DB 2761 GTGGGCAAGGAGCGGCTGTGCGGCGGAGGAGCTGTGCACTTGTGACCGCAAGCCCGCAG 2820
QY 2821 GAGCTGAGCTTCTGTGCTGTGAGGCGCTGTGGA 2853
DB 2821 GAGCTGAGCTTCTGTGCTGTGAGGCGCTGTGGA 2853

```

RESULT 2
 US-09-965-631-7
 : Sequence 7, Application US/09965631
 : Patent No. US20020115842A1
 : GENERAL INFORMATION:
 : APPLICANT: Fridge, Carl Johan
 : APPLICANT: Hilbun, Erin
 : TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Enco
 : FILE REFERENCE: LEX-0241-USA

CURRENT APPLICATION NUMBER: US/09/965,631
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/236,689
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 3446
TYPE: DNA
ORGANISM: homo sapiens
US-09-965-631-7

Query Match 100.0%; Score 2853; DB 10; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTCTAACCTGCTTTCGCGGGCGAACCCTGAGGCTCTGAG 60
DB ATGCTTCTGCTGGGATCTCTAACCTGCTTTCGCGGGCGAACCCTGAGGCTCTGAG 456
QY 61 CCAGAGCGGGAGGTAGTCTGCTCCATCCGACGTGACCCGGACATTAAGCGCGCGCTAC 120
DB CCAGAGCGGGAGGTAGTCTGCTCCATCCGACGTGACCCGGACATTAAGCGCGCGCTAC 516
QY 457 CCAGAGCGGGAGGTAGTCTGCTCCATCCGACGTGACCCGGACATTAAGCGCGCGCTAC 516
DB CCAGAGCGGGAGGTAGTCTGCTCCATCCGACGTGACCCGGACATTAAGCGCGCGCTAC 516
QY 121 TACTGGCGGGGTCCTCCGAGACTCCGGGGATCAGGAGCTATTTTCAGATCAAGCATTT 180
DB TACTGGCGGGGTCCTCCGAGACTCCGGGGATCAGGAGCTATTTTCAGATCAAGCATTT 576
QY 517 TACTGGCGGGGTCCTCCGAGACTCCGGGGATCAGGAGCTATTTTCAGATCAAGCATTT 576
DB TACTGGCGGGGTCCTCCGAGACTCCGGGGATCAGGAGCTATTTTCAGATCAAGCATTT 576
QY 181 CAGGAGACTTTTACTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 240
DB CAGGAGACTTTTACTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 636
QY 577 CAGGAGACTTTTACTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 636
DB CAGGAGACTTTTACTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 636
QY 241 ACTGAGCATGTGGGCGTCCCGCTCCAGAGGCTCAGCGGGGCTTTCAGACCTGCGACGC 300
DB ACTGAGCATGTGGGCGTCCCGCTCCAGAGGCTCAGCGGGGCTTTCAGACCTGCGACGC 696
QY 637 ACTGAGCATGTGGGCGTCCCGCTCCAGAGGCTCAGCGGGGCTTTCAGACCTGCGACGC 696
DB ACTGAGCATGTGGGCGTCCCGCTCCAGAGGCTCAGCGGGGCTTTCAGACCTGCGACGC 696
QY 301 TGCCTTATTTCTGGGAGCTGAACGCCGAGCCGACTGTTGCTGCTGCTGAGCCTGTGC 360
DB TGCCTTATTTCTGGGAGCTGAACGCCGAGCCGACTGTTGCTGCTGCTGAGCCTGTGC 756
QY 697 TGCCTTATTTCTGGGAGCTGAACGCCGAGCCGACTGTTGCTGCTGCTGAGCCTGTGC 756
DB TGCCTTATTTCTGGGAGCTGAACGCCGAGCCGACTGTTGCTGCTGCTGAGCCTGTGC 756
QY 361 GGGGGGCTCCGGGAGCTTGGCTACCGAGGGCGGAGATGTCATTAAGCCGCTGCC 420
DB GGGGGGCTCCGGGAGCTTGGCTACCGAGGGCGGAGATGTCATTAAGCCGCTGCC 816
QY 757 GGGGGGCTCCGGGAGCTTGGCTACCGAGGGCGGAGATGTCATTAAGCCGCTGCC 816
DB GGGGGGCTCCGGGAGCTTGGCTACCGAGGGCGGAGATGTCATTAAGCCGCTGCC 816
QY 421 AATGCTAGCGCGCGGCGGAGCGACAGCAGCAGGAGGCGACACCTTCTCCAGCCCGG 480
DB AATGCTAGCGCGCGGCGGAGCGACAGCAGCAGGAGGCGACACCTTCTCCAGCCCGG 876
QY 817 AATGCTAGCGCGCGGCGGAGCGACAGCAGCAGGAGGCGACACCTTCTCCAGCCCGG 876
DB AATGCTAGCGCGCGGCGGAGCGACAGCAGCAGGAGGCGACACCTTCTCCAGCCCGG 876
QY 481 GGTGTTCCGGGGGCTTCCGAGACCCCACTCTCGCTGCGGGGTGGCTCTGGGCTGG 540
DB GGTGTTCCGGGGGCTTCCGAGACCCCACTCTCGCTGCGGGGTGGCTCTGGGCTGG 936
QY 877 GGTGTTCCGGGGGCTTCCGAGACCCCACTCTCGCTGCGGGGTGGCTCTGGGCTGG 936
DB GGTGTTCCGGGGGCTTCCGAGACCCCACTCTCGCTGCGGGGTGGCTCTGGGCTGG 936
QY 541 AATGCTAGCGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 600
DB AATGCTAGCGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 996
QY 937 AATGCTAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 996
DB AATGCTAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 996
QY 601 AGTGTAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 660
DB AGTGTAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1056
QY 997 AGTGTAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1056
DB AGTGTAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1056
QY 661 GAGAGCGTGTGTGCTGCGGAGCGAGTCAAGTGTCAAGCGCGGAGCGGAGCGGAGCAT 720
DB GAGAGCGTGTGTGCTGCGGAGCGAGTCAAGTGTCAAGCGCGGAGCGGAGCGGAGCAT 1116
QY 1057 GAGAGCGTGTGTGCTGCGGAGCGAGTCAAGTGTCAAGCGCGGAGCGGAGCGGAGCAT 1116
DB GAGAGCGTGTGTGCTGCGGAGCGAGTCAAGTGTCAAGCGCGGAGCGGAGCGGAGCAT 1116
QY 721 TATCTGCTGAGCGCTGCTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCAT 780
DB TATCTGCTGAGCGCTGCTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCAT 1176
QY 1117 TATCTGCTGAGCGCTGCTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCAT 1176
DB TATCTGCTGAGCGCTGCTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCAT 1176
QY 781 CCCATCAACATCTGTTGGTCAAGGTGCTGCTTCTAGAGATCTGACTCCGGGCGCAAG 840
DB CCCATCAACATCTGTTGGTCAAGGTGCTGCTTCTAGAGATCTGACTCCGGGCGCAAG 1236
QY 1177 CCCATCAACATCTGTTGGTCAAGGTGCTGCTTCTAGAGATCTGACTCCGGGCGCAAG 1236
DB CCCATCAACATCTGTTGGTCAAGGTGCTGCTTCTAGAGATCTGACTCCGGGCGCAAG 1236
QY 841 GTACCGGCAATGCGCGCTGACGCTGCGCAACTTCTGCTGCGGAGGAGAAAGCTGAC 900
DB GTACCGGCAATGCGCGCTGACGCTGCGCAACTTCTGCTGCGGAGGAGAAAGCTGAC 900

DB 1237 GTACCGGCAATGCGCGCTGACGCTGCGCAACTTCTGCTGCGGAGGAGAAAGCTGAC 1296
QY 901 AATGAGTGAACAGCAACCCGAGTACTGGGAGCACTGCACTTCTTCCAGCAGGAGAG 960
DB AATGAGTGAACAGCAACCCGAGTACTGGGAGCACTGCACTTCTTCCAGCAGGAGAG 1356
QY 961 CTGTGTGAGGCGCACCACTGTGACACCCCTGGGAGCTGAGTGTGAGTGTGAG 1020
DB CTGTGTGAGGCGCACCACTGTGACACCCCTGGGAGCTGAGTGTGAGTGTGAG 1416
QY 1357 CTGTGTGAGGCGCACCACTGTGACACCCCTGGGAGCTGAGTGTGAGTGTGAG 1416
DB CTGTGTGAGGCGCACCACTGTGACACCCCTGGGAGCTGAGTGTGAGTGTGAG 1416
QY 1021 CCAGAGAGAGCTGCTGCTGATTTAGAGAGCATGGGCTTCCATAGCTTCCACCTGCG 1080
DB CCAGAGAGAGCTGCTGCTGATTTAGAGAGCATGGGCTTCCATAGCTTCCACCTGCG 1476
QY 1417 CCAGAGAGAGCTGCTGCTGATTTAGAGAGCATGGGCTTCCATAGCTTCCACCTGCG 1476
DB CCAGAGAGAGCTGCTGCTGATTTAGAGAGCATGGGCTTCCATAGCTTCCACCTGCG 1476
QY 1081 CACGAGCTGGGCGACGTTGCAATGATGCGCCATGACATGTAAGTGTGAGAGAGTGT 1140
DB CACGAGCTGGGCGACGTTGCAATGATGCGCCATGACATGTAAGTGTGAGAGAGTGT 1536
QY 1477 CACGAGCTGGGCGACGTTGCAATGATGCGCCATGACATGTAAGTGTGAGAGAGTGT 1536
DB CACGAGCTGGGCGACGTTGCAATGATGCGCCATGACATGTAAGTGTGAGAGAGTGT 1536
QY 1141 TTTGGGAAGCTCCGAGCCACACATGATGTCCCGAGCCCTCATCAGATGACCGTGGC 1200
DB TTTGGGAAGCTCCGAGCCACACATGATGTCCCGAGCCCTCATCAGATGACCGTGGC 1596
QY 1537 TTTGGGAAGCTCCGAGCCACACATGATGTCCCGAGCCCTCATCAGATGACCGTGGC 1596
DB TTTGGGAAGCTCCGAGCCACACATGATGTCCCGAGCCCTCATCAGATGACCGTGGC 1596
QY 1201 AATCCCTGTGAGCTGCTGAGTGTGCTGCTGATTCATCAGCAGCTTCTGAGAGCGGAG 1260
DB AATCCCTGTGAGCTGCTGAGTGTGCTGCTGATTCATCAGCAGCTTCTGAGAGCGGAG 1656
QY 1597 AATCCCTGTGAGCTGCTGAGTGTGCTGCTGATTCATCAGCAGCTTCTGAGAGCGGAG 1656
DB AATCCCTGTGAGCTGCTGAGTGTGCTGCTGATTCATCAGCAGCTTCTGAGAGCGGAG 1656
QY 1261 GACTGCTGCTGAGAGCAACCCAGCAAGCCATGCTCCCGGAGAGATCTCCCGGCGCC 1320
DB GACTGCTGCTGAGAGCAACCCAGCAAGCCATGCTCCCGGAGAGATCTCCCGGCGCC 1716
QY 1657 GACTGCTGCTGAGAGCAACCCAGCAAGCCATGCTCCCGGAGAGATCTCCCGGCGCC 1716
DB GACTGCTGCTGAGAGCAACCCAGCAAGCCATGCTCCCGGAGAGATCTCCCGGCGCC 1716
QY 1321 AGCTACACCTGAGGCGACAGTGTGAGTGTGCTTGGCTGGGCTCCAGCCTGTGCT 1380
DB AGCTACACCTGAGGCGACAGTGTGAGTGTGCTTGGCTGGGCTCCAGCCTGTGCT 1776
QY 1717 AGCTACACCTGAGGCGACAGTGTGAGTGTGCTTGGCTGGGCTCCAGCCTGTGCT 1776
DB AGCTACACCTGAGGCGACAGTGTGAGTGTGCTTGGCTGGGCTCCAGCCTGTGCT 1776
QY 1381 TACATGACATGATGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1440
DB TACATGACATGATGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1836
QY 1777 TACATGACATGATGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1836
DB TACATGACATGATGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1836
QY 1441 CAGAGCGGCACTTCCCTGCGGAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1500
DB CAGAGCGGCACTTCCCTGCGGAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1896
QY 1837 CAGAGCGGCACTTCCCTGCGGAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1896
DB CAGAGCGGCACTTCCCTGCGGAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1896
QY 1501 AAGGGGCTGCTGAGGAGAGACACAACCTCAACAGCAAGAGGTGATGTTCTGCGGCG 1560
DB AAGGGGCTGCTGAGGAGAGAGACACAACCTCAACAGCAAGAGGTGATGTTCTGCGGCG 1956
QY 1897 AAGGGGCTGCTGAGGAGAGAGACACAACCTCAACAGCAAGAGGTGATGTTCTGCGGCG 1956
DB AAGGGGCTGCTGAGGAGAGAGAGACACAACCTCAACAGCAAGAGGTGATGTTCTGCGGCG 1956
QY 1561 AATGGGATCCTATGAGGCTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1620
DB AATGGGATCCTATGAGGCTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2016
QY 1957 AATGGGATCCTATGAGGCTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2016
DB AATGGGATCCTATGAGGCTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2016
QY 1621 AGGAGTGTGACCAACCCAGCCTGCGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGCTG 1680
DB AGGAGTGTGACCAACCCAGCCTGCGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGCTG 2076
QY 2017 AGGAGTGTGACCAACCCAGCCTGCGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGCTG 2076
DB AGGAGTGTGACCAACCCAGCCTGCGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGCTG 2076
QY 1681 AATATCCGATCCTGCAATGTGAGAGCCTGCGGAGCTGAGCTGCGGAGAAAGCTTCCGG 1740
DB AATATCCGATCCTGCAATGTGAGAGCCTGCGGAGCTGAGCTGCGGAGAAAGCTTCCGG 2136
QY 2077 AATATCCGATCCTGCAATGTGAGAGCCTGCGGAGCTGAGCTGCGGAGAAAGCTTCCGG 2136
DB AATATCCGATCCTGCAATGTGAGAGCCTGCGGAGCTGAGCTGCGGAGAAAGCTTCCGG 2136
QY 1741 GAGGAGCACTGTGAGGCTTCAAGGCTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1800
DB GAGGAGCACTGTGAGGCTTCAAGGCTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2196
QY 2137 GAGGAGCACTGTGAGGCTTCAAGGCTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2196
DB GAGGAGCACTGTGAGGCTTCAAGGCTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2196
QY 1801 GTGGGCACTGGGCTCCCAAGTACTCCGCGCTGCTCCCGGAGCAAGTCAAGCTCATCTGC 1860
DB GTGGGCACTGGGCTCCCAAGTACTCCGCGCTGCTCCCGGAGCAAGTCAAGCTCATCTGC 2256
QY 2197 GTGGGCACTGGGCTCCCAAGTACTCCGCGCTGCTCCCGGAGCAAGTCAAGCTCATCTGC 2256
DB GTGGGCACTGGGCTCCCAAGTACTCCGCGCTGCTCCCGGAGCAAGTCAAGCTCATCTGC 2256
QY 1861 CGAGGCAATGAGCACTGCTGCTGATTTATGCTGCGCAAGCAAGTGTGAGAGGAGAGCTG 1920
DB CGAGGCAATGAGCACTGCTGCTGATTTATGCTGCGCAAGCAAGTGTGAGAGGAGAGCTG 2316
QY 2257 CGAGGCAATGAGCACTGCTGCTGATTTATGCTGCGCAAGCAAGTGTGAGAGGAGAGCTG 2316
DB CGAGGCAATGAGCACTGCTGCTGATTTATGCTGCGCAAGCAAGTGTGAGAGGAGAGCTG 2316
QY 1921 TGTCTCTGAGCTCCACGCTCGGCTGTGTCGAAGGCAAGTGAAGGCTGCTGAT 1980
DB TGTCTCTGAGCTCCACGCTCGGCTGTGTCGAAGGCAAGTGAAGGCTGCTGAT 1980


```
Db 2317 TGCTCTCTGATCCACCTCCGCTGTGTGTCCAGGCAAGTGCATCAAGCTGGCTGTAT 2376
QY 1981 GGGAACTGGGCTCCAAAGAGATTGCAAGTGTGGGCTGTGGGGGAGCAATAAG 2040
Db 2377 GGGAACTGGGCTCCAAAGAGATTGCAAGTGTGGGCTGTGGGGGAGCAATAAG 2436
QY 2041 AGCTGCAAGAGTGTGACTGTGACTCTTCAACCAAGCCCATGTGCTACAAATTCGTGTG 2100
Db 2437 AGCTGCAAGAGTGTGACTGTGACTCTTCAACCAAGCCCATGTGCTACAAATTCGTGTG 2496
QY 2101 GGCATCCCCCGAGGGGCTCAAGCATGACATCCGCCAGGCGGTTCACAAAGGCTGATC 2160
Db 2497 GGCATCCCCCGAGGGGCTCAAGCATGACATCCGCCAGGCGGTTCACAAAGGCTGATC 2556
QY 2161 GGGGATGCAACTACTGTGCTGTGAAGACAGCCCAAGTACTGCTCAACGGGCAAT 2220
Db 2557 GGGGATGCAACTACTGTGCTGTGAAGACAGCCCAAGTACTGCTCAACGGGCAAT 2516
QY 2221 TTGCTGTGTGTGGGCTGTGGAGCGGGACCTGTGTGTGAAGGCGAGTCTGTGGGTACAGC 2280
Db 2617 TTGCTGTGTGTGGGCTGTGGAGCGGGACCTGTGTGTGAAGGCGAGTCTGTGGGTACAGC 2676
QY 2281 GGCACGGGACACAGCGGTGTGAAGAGCTTCAGAGCTCCCGGCCCATCTGAGCGGCTGAGC 2340
Db 2677 GGCACGGGACACAGCGGTGTGAAGAGCTTCAGAGCTCCCGGCCCATCTGAGCGGCTGAGC 2736
QY 2341 GTGAGAGTCTCTCCGTGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2737 GTGAGAGTCTCTCCGTGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAAGAGCTGTGGGGAAGATGCTCTCATTCCTCAAGAGACCCCGGGGAGCCCTCTGTG 2460
Db 2797 CCCAAAGAGCTGTGGGGAAGATGCTCTCATTCCTCAAGAGACCCCGGGGAGCCCTCTGTG 2856
QY 2461 TTGCACAACAGCGTCTCTACGCTCTCCAAACAGGTGGAGAGCGCGGACAGACGCCCTT 2520
Db 2857 TTGCACAACAGCGTCTCTACGCTCTCCAAACAGGTGGAGAGCGCGGACAGACGCCCTT 2916
QY 2521 GCACGCTGGGTGCTGTGGCACTGTGGGGCGCTGCTCCGCGAGCTGTGGGCACTGTGCA 2580
Db 2917 GCACGCTGGGTGCTGTGGCACTGTGGGGCGCTGCTCCGCGAGCTGTGGGCACTGTGCA 2976
QY 2581 AAGCGGCGGTGTGACTGTGCGGGCTCGCGGCGAGCGGAGCGGCTCTCTGTGTATGA 2640
Db 2977 AAGCGGCGGTGTGACTGTGCGGGCTCGCGGCGAGCGGAGCGGCTCTCTGTGTATGA 3036
QY 2641 GCCCATCGGCGGTGTGAGACAAAGCTGTGGGGAGCGGCTGTGCCCACTGTGGAGCTCAGC 2700
Db 3037 GCCCATCGGCGGTGTGAGACAAAGCTGTGGGGAGCGGCTGTGCCCACTGTGGAGCTCAGC 3096
QY 2701 GCGTGTGACACCTGTGCTCCAAAGACTGTGGGCGGCGGGAATTCAGAGCGGCTCAAGTGT 2760
Db 3097 GCGTGTGACACCTGTGCTCCAAAGACTGTGGGCGGCGGGAATTCAGAGCGGCTCAAGTGT 3156
QY 2761 GTGGGCGACAGAGCGGCGGTGTGGGCGGCGGAGACAGTGTGACAGCGGCAAGGCCCA 2820
Db 3157 GTGGGCGACAGAGCGGCGGTGTGGGCGGCGGAGACAGTGTGACAGCGGCAAGGCCCA 3216
QY 2821 GAGCTGACTTCTGCTCTGAGGCGGCTGTGA 2853
Db 3217 GAGCTGACTTCTGCTCTGAGGCGGCTGTGA 3249
```

RESULT 3

```
US-10-163-316-3
; Sequence 3, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: Therefore
; FILE REFERENCE: MP101-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
```

```
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-10-163-316-3

Query Match          52.3%; Score 1491; DB 9; Length 2469;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTTAACCTTGCTTCCGCGGAGACCGGTGAGGCTGTGAG 60
Db 1 ATGCTTCTGCTGGGATCTTAACCTTGCTTCCGCGGAGACCGGTGAGGCTGTGAG 60
QY 61 CCAGAGCGGAGGTAGTGTTCCTCAATCCAGCTGACCCCGGACATTACGGCGGCTAC 120
Db 61 CCAGAGCGGAGGTAGTGTTCCTCAATCCAGCTGACCCCGGACATTACGGCGGCTAC 120
QY 121 TACTGGGCGGAGTCCGAGAGACTCCGGGAGTCAAGGACTATTTTCAATACAGCAATT 180
Db 121 TACTGGGCGGAGTCCGAGAGACTCCGGGAGTCAAGGACTATTTTCAATACAGCAATT 180
QY 181 CAGAGAGACTTTTACCTACACTGACCGCGGATGCTAGTTCCTTGGCTCCGCTTCC 240
Db 181 CAGAGAGACTTTTACCTACACTGACCGCGGATGCTAGTTCCTTGGCTCCGCTTCC 240
QY 241 ACTGAGCATGTGGGGTCTCCCTCCAGAGGCTCACCGGGGCTTTAGACTTGCAGC 300
Db 241 ACTGAGCATGTGGGGTCTCCCTCCAGAGGCTCACCGGGGCTTTAGACTTGCAGC 300
QY 301 TGCTTCTATTTCTGGGAGGTGAGACCGGAGCGGAGTGTGCTGCTGAGCTGTG 360
Db 301 TGCTTCTATTTCTGGGAGGTGAGACCGGAGCGGAGTGTGCTGCTGAGCTGTG 360
QY 361 GGGGGGCTTCCGCGGAGCTTTGGCTTACCGAGCGCGGAGTATGTATTAAGCCGCTCC 420
Db 361 GGGGGGCTTCCGCGGAGCTTTGGCTTACCGAGCGCGGAGTATGTATTAAGCCGCTCC 420
QY 421 AATGTACGCGGCGGCGGCGGAGCAAGCCAGAGCGGAGCTTTCACAGCGCGG 480
Db 421 AATGTACGCGGCGGCGGCGGAGCAAGCCAGAGCGGAGCTTTCACAGCGCGG 480
QY 481 GGTGTTCGCGGCGGCTTCCGAGACCCCACTCTGTGCGGGGTGGCTCGGGTGG 540
Db 481 GGTGTTCGCGGCGGCTTCCGAGACCCCACTCTGTGCGGGGTGGCTCGGGTGG 540
QY 541 AACCCCGCATCTTACAGGCGCTTGAGCCCTTACAGCCCGGGGGGCTTGGGGAG 600
Db 541 AACCCCGCATCTTACAGGCGCTTGAGCCCTTACAGCCCGGGGGGCTTGGGGAG 600
QY 601 AGTGTACCGCGGCGGAGTGTGGGCGGCGCAAGGCTTGTGTATCCGCGGTTAC 660
Db 601 AGTGTACCGCGGCGGAGTGTGGGCGGCGCAAGGCTTGTGTATCCGCGGTTAC 660
QY 661 GAGAGCGTGTGTGCGGAGAGTCAATGATTCACAGGCGGAGCACTGTGAACAT 720
Db 661 GAGAGCGTGTGTGCGGAGAGTCAATGATTCACAGGCGGAGCACTGTGAACAT 720
QY 721 TATGTCTGAGAGCTGTGCAAGCGGCGGAGCTTACCGCATCCAGACTCTCAAC 780
Db 721 TATGTCTGAGAGCTGTGCAAGCGGCGGAGCTTACCGCATCCAGACTCTCAAC 780
QY 781 CCCATCAACATCGTTGTGTCAAGGTGCTGCTTCTTGAAGATGTGACTCCGGGCCAAG 840
Db 781 CCCATCAACATCGTTGTGTCAAGGTGCTGCTTCTTGAAGATGTGACTCCGGGCCAAG 840
```

```

QY 841 GTACCGGCAATGCGGCGCTGACGGTGGCACTTCTGTGCTGGGAGAGAGCT3AAC 900
Db 841 GTACCGGCAATGCGGCGCTGACGGTGGCACTTCTGTGCTGGGAGAGAGCT3AAC 900
QY 901 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCCATCTCTTCAACGAGGAC 960
Db 901 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCCATCTCTTCAACGAGGAC 960
QY 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGGATGGCTGATGTGGTACCTGTGTAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGGATGGCTGATGTGGTACCTGTGTAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGTCAATGTAGAGAGATGGGCTTCCATCAGCCTTCAACACTGCC 1080
Db 1021 CCCAAGAGAGCTGCTCTGTCAATGTAGAGAGATGGGCTTCCATCAGCCTTCAACACTGCC 1080
QY 1081 CACGAGCTGGGCGACGTGTTCACATGCCCCATGACATGTGAAGTGTGTAGAGAGGTG 1140
Db 1081 CACGAGCTGGGCGACGTGTTCACATGCCCCATGACATGTGAAGTGTGTAGAGAGGTG 1140
QY 1141 TTGTGGAGACTCGGACCAACCATGATGTCCCGACCTCATCCAGATCCAGCTGCG 1200
Db 1141 TTGTGGAGACTCGGACCAACCATGATGTCCCGACCTCATCCAGATCCAGCTGCG 1200
QY 1201 AACCCCTGGTCACTGTGACAGTGTGCTGATCATCAGCCTTCTTGAAGCGGCGACGT 1260
Db 1201 AACCCCTGGTCACTGTGACAGTGTGCTGATCATCAGCCTTCTTGAAGCGGCGACGT 1260
QY 1261 GACTGCTCTCTGGACCAACCCAGACGCCCATCTCCCTCCGAGAGATCTGCGCGGCGCC 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGACGCCCATCTCCCTCCGAGAGATCTGCGCGGCGCC 1320
QY 1321 AGCTAACCTGTAGCCAGACAGTGTGAGTGGCTTTTGGGCTGGCTCCAAAGCCCTGTCT 1380
Db 1321 AGCTAACCTGTAGCCAGACAGTGTGAGTGGCTTTTGGGCTGGCTCCAAAGCCCTGTCT 1380
QY 1381 TATCATCAGTACTGACACCAAGTGTGTGACACGGGAGCCAAAGAGATGTGTCTC 1440
Db 1381 TATCATCAGTACTGACACCAAGTGTGTGACACGGGAGCCAAAGAGATGTGTCTC 1440
QY 1441 CAGACCGGCACTTCCCTCTGGGCGGATGGACACAGCTGTGGCGAGGCAAGCTTGGCTTC 1500
Db 1441 CAGACCGGCACTTCCCTCTGGGCGGATGGACACAGCTGTGGCGAGGCAAGCTTGGCTTC 1500
QY 1501 AAAGGGGCTGCTGGAGAGACACACTTCAACAGCAGG 1542
Db 1501 AAAGGGGCTGCTGGAGAGACACACTTCAACAGCAGG 1542

RESULT 4
US-10-163-316-1
; Sequence 1, Application US/10163316
; Publication No. US2002019703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025PIRM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(2940)
; OTHER INFORMATION: n = A,T,C or G
US-10-163-316-1

Query Match 52.3%; Score 1491; DB 9; Length 2940;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCAATCCCTAACCTGCTTTCGCGGGGGAACCGGTGAGGCTCTGAG 60
Db 472 ATGCTTCTGCTGGGCAATCCCTAACCTGCTTTCGCGGGGGAACCGGTGAGGCTCTGAG 531
QY 61 CCAGAGCGGAGAGTACGTCTTCCATCCGATCGATGAGACCCGAGCATTTAAAGCGCGGCTAC 120
Db 532 CCAGAGCGGAGAGTACGTCTTCCATCCGATCGATGAGACCCGAGCATTTAAAGCGCGGCTAC 591
QY 121 TACTGGCGGGGTCCCGAGAGACTCCGGGGATCAGGAGACTATTTTTCAGATCAGCAAT 180
Db 592 TACTGGCGGGGTCCCGAGAGACTCCGGGGATCAGGAGACTATTTTTCAGATCAGCAAT 651
QY 181 CAGAGAGACTTTTACCTACCTGACGCGGAGATGCTAGTCTTGGGCTCCGCGCTTCTCC 240
Db 652 CAGAGAGACTTTTACCTACCTGACGCGGAGATGCTAGTCTTGGGCTCCGCGCTTCTCC 711
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGAGGCTCACCGGGGGCTCTTGAAGACTGCGACAC 300
Db 712 ACTGAGCATCTGGGCGTCCCTCCAGAGGCTCACCGGGGGCTCTTGAAGACTGCGACAC 771
QY 301 TGTCTTCTATCTGGGGAGCTGACAGCCGAGCCGAGCTGTGCTGCTGTGAGCTGTGCT 360
Db 772 TGTCTTCTATCTGGGGAGCTGACAGCCGAGCCGAGCTGTGCTGCTGTGAGCTGTGCT 831
QY 361 GGGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCGGATGTGATTAAGCGCGCTGCC 420
Db 832 GGGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCGGATGTGATTAAGCGCGCTGCC 891
QY 421 AATGTACAGCGCGCGCGCGCGAGCGCAACAGCCAGGCGGAGCAACCTTCTTCAAGCGCG 480
Db 892 AATGTACAGCGCGCGCGCGCGAGCGCAACAGCCAGGCGGAGCAACCTTCTTCAAGCGCG 951
QY 481 GGTGTTCGGGCGGCGCTTCCGAGAGCCCACTCTGCTCGCGGGGTGGCTCGGCTGCG 540
Db 952 GGTGTTCGGGCGGCGCTTCCGAGAGCCCACTCTGCTCGCGGGGTGGCTCGGCTGCG 1011
QY 541 AACCCCGCATCTCTGCGGGGCGCTGAGCCCTTAAAGCGCGGCGGCGGCGCTTGGGAG 600
Db 1012 AACCCCGCATCTCTGCGGGGCGCTGAGCCCTTAAAGCGCGGCGGCGGCGCTTGGGAG 1071
QY 601 AGTCGTAGCGCGCGAGGCTGTGGGCGCGCAAGCGCTTTCGTGTATCCCGGGTACGTG 660
Db 1072 AGTCGTAGCGCGCGAGGCTGTGGGCGCGCAAGCGCTTTCGTGTATCCCGGGTACGTG 1131
QY 661 GAGACGCTGTGTGTGCGCGAGAGTCAATGTCAAGTTTCAAGCGCGGAGACTTGGACAT 720
Db 1132 GAGACGCTGTGTGTGCGCGAGAGTCAATGTCAAGTTTCAAGCGCGGAGACTTGGACAT 1191
QY 721 TATCTGTGAGCGCTGTGGCAAGGCGGCGGAGCTCTACCGCATCCAGATCCCTCAAC 780
Db 1192 TATCTGTGAGCGCTGTGGCAAGGCGGCGGAGCTCTACCGCATCCAGATCCCTCAAC 1251
QY 781 CCCATCAACATCGTTGTGTCAAGGTGTGCTTCTTAAAGATGTGACTCCGGGCCCAAG 840
Db 1252 CCCATCAACATCGTTGTGTCAAGGTGTGCTTCTTAAAGATGTGACTCCGGGCCCAAG 1311
QY 841 GTACCGGCAATGCGGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAAAGAGTGTAC 900
Db 1312 GTACCGGCAATGCGGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAAAGAGTGTAC 1371
QY 901 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCACTGCACTCTTCAACAGGAGCAG 960
Db 1372 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCACTGCACTCTTCAACAGGAGCAG 1431

```

QY	121	TACTGCGGGGTC	CCGAGAGACTCCG	GGGATCAGG	AGACTATTTTCAGT	CACAGCATTT	180
Db	121	TACTGCGGGGTC	CCGAGAGACTCCG	GGGATCAGG	AGACTATTTTCAGT	CACAGCATTT	180
QY	181	CAGGAGAGACTTTT	ACCTTACACCTG	ACGCGGATGCT	CTAGTTCTTG	GGCTCCGCTTCTCC	240
Db	181	CAGGAGAGACTTTT	ACCTTACACCTG	ACGCGGATGCT	CTAGTTCTTG	GGCTCCGCTTCTCC	240
QY	241	ACTGAGCATCTT	GGGGGTGTC	CCCCCTC	CAAGGGGTC	ACACCGGGGCTTTCAGACCTTGGACACG	300
Db	241	ACTGAGCATCTT	GGGGGTGTC	CCCCCTC	CAAGGGGTC	ACACCGGGGCTTTCAGACCTTGGACACG	300
QY	301	TGCTTTATTTCTG	GGGAGAGCTGA	ACGCGAGCTG	TTCGCTGTG	TGAGACCTGTGC	360
Db	301	TGCTTTATTTCTG	GGGAGAGCTGA	ACGCGAGCTG	TTCGCTGTG	TGAGACCTGTGC	360
QY	361	GGGGGGCTTC	CGGAGACTTTGG	CTTGGCTAC	GAGGCGCGATATGT	CATTAGCCGCTGCC	420
Db	361	GGGGGGCTTC	CGGAGACTTTGG	CTTGGCTAC	GAGGCGCGATATGT	CATTAGCCGCTGCC	420
QY	421	AATGCTTACGCGG	CGGCGGCGAC	AGCAACGAC	AGGGGCGCACACCTTCTCC	ACGCGCGG	480
Db	421	AATGCTTACGCGG	CGGCGGCGAC	AGCAACGAC	AGGGGCGCACACCTTCTCC	ACGCGCGG	480
QY	481	GGTGTTCG	GGGCGGGCTTCC	CGGAGAGACCC	ACCTTCTGCTG	CGGGGGTGGCTTGGGCTTG	540
Db	481	GGTGTTCG	GGGCGGGCTTCC	CGGAGAGACCC	ACCTTCTGCTG	CGGGGGTGGCTTGGGCTTG	540
QY	541	AAACCCGCGCAT	CTTACGAGGCTTGG	ACCTTTCACA	CGCGCGGCGGGGCTTCC	GGGGAG	600
Db	541	AAACCCGCGCAT	CTTACGAGGCTTGG	ACCTTTCACA	CGCGCGGCGGGGCTTCC	GGGGAG	600
QY	601	AGTCGTACCGG	CGCAGAGTCTGGG	CGCGCCAA	GGGTTTGTGTCTAT	CCCGGGGTACG	660
Db	601	AGTCGTACCGG	CGCAGAGTCTGGG	CGCGCCAA	GGGTTTGTGTCTAT	CCCGGGGTACG	660
QY	661	GAGAGCTGTGTG	TGCGCGGAGCA	GTCAATGTC	CAAGTTTCAC	GCGCGCGACCTGTGAACAT	720
Db	661	GAGAGCTGTGTG	TGCGCGGAGCA	GTCAATGTC	CAAGTTTCAC	GCGCGCGACCTGTGAACAT	720
QY	721	TATCTGCTGAG	CGTGTGTCGACA	AGGTGCTCTT	AGAGATGTG	ATCCTCGGGGCCAAG	780
Db	721	TATCTGCTGAG	CGTGTGTCGACA	AGGTGCTCTT	AGAGATGTG	ATCCTCGGGGCCAAG	780
QY	781	CCCATCAACAT	CGTTGTGTGTC	CAAGGTGCTCTT	AGAGATGTG	ATCCTCGGGGCCAAG	840
Db	781	CCCATCAACAT	CGTTGTGTGTC	CAAGGTGCTCTT	AGAGATGTG	ATCCTCGGGGCCAAG	840
QY	841	GTCACCGGCA	TGCGGGCTTG	AGCGCTGCGCA	CTTGTGTGCTG	CGCGCAGAAGAAAGCTGAAC	900
Db	841	GTCACCGGCA	TGCGGGCTTG	AGCGCTGCGCA	CTTGTGTGCTG	CGCGCAGAAGAAAGCTGAAC	900
QY	901	AAATGTAGTGA	CAAGCAACCCCG	AGTACTGTGG	ACACCTCCATCTT	CCACGAGGACAG	960
Db	901	AAATGTAGTGA	CAAGCAACCCCG	AGTACTGTGG	ACACCTCCATCTT	CCACGAGGACAG	960
QY	961	CTGTGTGAG	CGACACACCTGT	GAACCCCTGG	AGATGGGATGTGG	GTGATCCATGTGTAC	1020
Db	961	CTGTGTGAG	CGACACACCTGT	GAACCCCTGG	AGATGGGATGTGG	GTGATCCATGTGTAC	1020
QY	1021	CCCAAGAAG	AGCTGTCTGT	CTATTTGA	GGAGAGATGGG	CTTCCATCAACCTTCAACCTGCC	1080
Db	1021	CCCAAGAAG	AGCTGTCTGT	CTATTTGA	GGAGAGATGGG	CTTCCATCAACCTTCAACCTGCC	1080
QY	1081	CACGAGCTGG	6 1091				
Db	1081	CACGAGCTGG	6 1091				

```
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Fildale, Carl Johan
; APPLICANT: Halbur, Erin
; TITLE OF INVENTION: No. US20020115842A1el Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US-09/965,631
; PRIORITY FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 966
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-965-631-1

Query Match      33.6%; Score 959; DB 10; Length 966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCTTCTGCTGGGACATCTTAACCCCTTGTGGCCGGGGAACCCGTGAGCTCTGAG 60
      1 ATGCTTCTGCTGGGACATCTTAACCCCTTGTGGCCGGGGAACCCGTGAG 60
Db      1 ATGCTTCTGCTGGGACATCTTAACCCCTTGTGGCCGGGGAACCCGTGAG 60

QY      61 CCAGAGCGGAGGAGTAGTCTTCCCATCCGACTGAGCCCGGACATTAACGCGCGCGCTGAC 120
      61 CCAGAGCGGAGGAGTAGTCTTCCCATCCGACTGAGCCCGGACATTAACGCGCGCGCTGAC 120
Db      61 CCAGAGCGGAGGAGTAGTCTTCCCATCCGACTGAGCCCGGACATTAACGCGCGCGCTGAC 120

QY      121 TACTGGCGGGGTCGCCGAGACTCCGGGATCAAGGACTATTTTCAGATCACAGCATTT 180
      121 TACTGGCGGGGTCGCCGAGACTCCGGGATCAAGGACTATTTTCAGATCACAGCATTT 180
Db      121 TACTGGCGGGGTCGCCGAGACTCCGGGATCAAGGACTATTTTCAGATCACAGCATTT 180

QY      181 CAGGAGGACTTTTACCTACACCTGACGCCGGATGCTCAGTTCTGCTCCCGCTCTCC 240
      181 CAGGAGGACTTTTACCTACACCTGACGCCGGATGCTCAGTTCTGCTCCCGCTCTCC 240
Db      181 CAGGAGGACTTTTACCTACACCTGACGCCGGATGCTCAGTTCTGCTCCCGCTCTCC 240

QY      241 ACTGACATCTGGGCTCCCTCCCTCAAGGGGCTCACCGGGGGCTTTACAGACTGCGACGC 300
      241 ACTGACATCTGGGCTCCCTCCCTCAAGGGGCTCACCGGGGGCTTTACAGACTGCGACGC 300
Db      241 ACTGACATCTGGGCTCCCTCCCTCAAGGGGCTCACCGGGGGCTTTACAGACTGCGACGC 300

QY      301 TGCTTCTATTCTGGGAGCGTAACGCCGAGCCGGAAGCTGCTGCTGCTGTGAGCTGTGC 360
      301 TGCTTCTATTCTGGGAGCGTAACGCCGAGCCGGAAGCTGCTGCTGCTGTGAGCTGTGC 360
Db      301 TGCTTCTATTCTGGGAGCGTAACGCCGAGCCGGAAGCTGCTGCTGCTGTGAGCTGTGC 360

QY      361 GGGGGGCTCCGGGAGCTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCCGCTGCC 420
      361 GGGGGGCTCCGGGAGCTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCCGCTGCC 420
Db      361 GGGGGGCTCCGGGAGCTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCCGCTGCC 420

QY      421 AATGCTAGCGCGCGCGGCGCAAGCGCAACAGCCAGCGGCGCACACCTTCTCCAGCGCGCG 480
      421 AATGCTAGCGCGCGCGGCGCAAGCGCAACAGCGGCGCACACCTTCTCCAGCGCGCG 480
Db      421 AATGCTAGCGCGCGCGGCGCAAGCGCAACAGCGGCGCACACCTTCTCCAGCGCGCG 480

QY      481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 540
      481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 540
Db      481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 540

QY      541 AACCCCGGCATCTACGGGCGCTTGAGACCTTACAAAGCGCGGGGGGCTTTCGGGGAG 600
      541 AACCCCGGCATCTACGGGCGCTTGAGACCTTACAAAGCGCGGGGGGCTTTCGGGGAG 600
Db      541 AACCCCGGCATCTACGGGCGCTTGAGACCTTACAAAGCGCGGGGGGCTTTCGGGGAG 600

QY      601 AGTCGTAGCGCGCGAGGCTTGGCGCGCAAGCGCTTCTGTCGTATCCCGGCTACG 660
      601 AGTCGTAGCGCGCGAGGCTTGGCGCGCAAGCGCTTCTGTCGTATCCCGGCTACG 660
Db      601 AGTCGTAGCGCGCGAGGCTTGGCGCGCAAGCGCTTCTGTCGTATCCCGGCTACG 660

QY      661 GAGACGCTGCTGTCGCGGAGAGTCAATGTCAGATTCCAGGGCGGAGCTTGAACAT 720
      661 GAGACGCTGCTGTCGCGGAGAGTCAATGTCAGATTCCAGGGCGGAGCTTGAACAT 720
Db      661 GAGACGCTGCTGTCGCGGAGAGTCAATGTCAGATTCCAGGGCGGAGCTTGAACAT 720

QY      721 TATCTGTGACGCTGCTGGCAAGCGCGGCGGACTCTACCGGCATCCAGATCCTCAAC 780
      721 TATCTGTGACGCTGCTGGCAAGCGCGGCGGACTCTACCGGCATCCAGATCCTCAAC 780
Db      721 TATCTGTGACGCTGCTGGCAAGCGCGGCGGACTCTACCGGCATCCAGATCCTCAAC 780
```

```
QY      781 CCATCAACATGTTGTGTCGAAGTGCCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840
      781 CCATCAACATGTTGTGTCGAAGTGCCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840
Db      781 CCATCAACATGTTGTGTCGAAGTGCCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840

QY      841 GTACCGGCAATGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGAGCTGAAC 900
      841 GTACCGGCAATGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGAGCTGAAC 900
Db      841 GTACCGGCAATGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGAGCTGAAC 900

QY      901 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCCATCCTTTCACCAAGGACAGA 959
      901 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCCATCCTTTCACCAAGGACAGA 959
Db      901 AAAGTAGTGACAGACACCCGAGTACTGGGACACTGCCATCCTTTCACCAAGGACAGA 959
```

Search completed: May 16, 2003, 10:21:28
Job time : 4157 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 01:01:57 ; Search time 5576 Seconds
(without alignments)
12864.332 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgctctgcgtggcctcct.....gcgtctgagcgctgtcga 2853

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 24791104 seqs, 12571243825 residues

Word size : 24

Total number of hits satisfying chosen parameters: 123

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/US06000_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06001_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095E_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095F_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095G_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095H_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095I_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US095J_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US095K_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US095L_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US095M_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US095N_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US095O_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US095P_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US095Q_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US095R_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US095S_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US095T_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US095U_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US095V_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US095W_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US095X_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US095Y_COMB.seq:*

```

```

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	2853	100.0	US-09-965-631-3	Sequence 3, Appl1
2	2853	100.0	US-09-965-631-7	Sequence 7, Appl1
3	2700	94.6	US-09-741-151-1	Sequence 1, Appl1
4	2695	94.5	US-10-009-332-2	Sequence 2, Appl1
5	2196	77.0	US-60-216-821-9	Sequence 9, Appl1
6	1491	52.3	US-10-163-316-3	Sequence 3, Appl1
7	1491	52.3	US-60-297-863-3	Sequence 3, Appl1
8	1491	52.3	US-10-163-316-1	Sequence 1, Appl1
9	1491	52.3	US-60-297-863-1	Sequence 1, Appl1
10	1488	52.2	US-10-093-463-27	Sequence 27, Appl1
11	1261	44.2	US-10-093-463-29	Sequence 29, Appl1
12	1261	44.2	US-60-212-656-762	Sequence 762, App
13	1261	44.2	US-60-242-679-1614	Sequence 1814, Ap
14	1091	38.2	US-09-965-631-5	Sequence 5, Appl1
15	959	33.6	US-09-965-631-1	Sequence 1, Appl1
16	958	33.6	US-60-230-435-2152	Sequence 2152, Ap
17	958	33.6	US-60-206-028-17	Sequence 17, Appl1
18	958	33.6	US-60-206-028-18	Sequence 18, Appl1
19	958	33.6	US-60-242-679-304	Sequence 304, App
20	958	33.6	US-09-741-151-3	Sequence 3, Appl1
21	958	33.6	US-60-207-315-28	Sequence 28, Appl1

```

c 22 958 33.6 31627 65 US-60-212-656-232 Sequence 232, App
c 23 958 33.6 32768 64 US-60-207-315-29 Sequence 29, Appl
c 24 958 33.6 32768 64 US-60-208-020-14 Sequence 14, Appl
c 25 958 33.6 32768 64 US-60-209-043-17 Sequence 17, Appl
c 26 958 33.6 32768 64 US-60-209-043-18 Sequence 18, Appl
c 27 958 33.6 32768 64 US-60-209-043-19 Sequence 19, Appl
c 28 958 33.6 32768 65 US-60-212-656-233 Sequence 233, App
c 29 958 33.6 32768 67 US-60-230-435-158 Sequence 158, App
c 30 958 33.6 32768 67 US-60-230-435-159 Sequence 159, App
c 31 844 29.6 1829 22 US-09-575-003-17 Sequence 17, Appl
c 32 844 29.6 1829 23 US-09-609-059-12 Sequence 2, Appl
c 33 844 29.6 1829 36 US-09-975-545-17 Sequence 17, Appl
c 34 844 29.6 1829 39 US-10-071-241-2 Sequence 2, Appl
c 35 838 29.4 31959 64 US-60-200-382-27 Sequence 27, Appl
c 36 838 29.4 31959 64 US-60-200-382-28 Sequence 28, Appl
c 37 724 25.4 1875 64 US-60-207-213-37 Sequence 37, Appl
c 38 724 25.4 1875 65 US-60-213-845-8 Sequence 8, Appl
c 39 639 22.4 741 65 US-60-213-845-538 Sequence 538, Appl
c 40 611 21.4 662 62 US-60-188-162-89 Sequence 1688, Ap
c 41 611 21.4 662 62 US-60-188-162-89 Sequence 89, Appl
c 42 575 19.5 633 63 US-60-196-174-3578 Sequence 493, Ap
c 43 555 19.1 555 39 US-10-093-463-35 Sequence 35, Appl
c 44 544 18.6 530 61 US-60-173-469-149 Sequence 149, App
c 45 530 18.6 530 61 US-60-173-469-149 Sequence 149, App
```

ALIGNMENTS

```

RESULT 1
US-09-965-631-3
: Sequence 3, Application US/09965631
: GENERAL INFORMATION:
: APPLICANT: Fritdelle, Carl Johan
: APPLICANT: Hildun, Erin
: FILE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
: TITLE REFERENCE: Lex-0241-USA
: CURRENT APPLICATION NUMBER: US/09/965,631
: PRIOR FILING DATE: 2001-09-27
: PRIOR APPLICATION NUMBER: US 60/236,689
: PRIOR FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 2853
: TYPE: DNA
: ORGANISM: homo sapiens
US-09-965-631-3
```

```

Query Match 100.0%; Score 2853; DB 36; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ATGCTTTGCTGGGCACTCTACACCTTGGCTTGGCGGCGAACCCTGGAGGCTCTGAG 60
DB 1 ATGCTTTGCTGGGCACTCTACACCTTGGCTTGGCGGCGAACCCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGTAGTCGTTCCCATCCGACTGGAGCCCGGACATTAACGCGCGCTAC 120
DB 61 CCAGAGCGGAGTAGTCGTTCCCATCCGACTGGAGCCCGGACATTAACGCGCGCTAC 120
QY 121 TACTGGGGGGTCCCGAGAGCTCCGAGGATCAGGAGCTCATTTTTCAGATCAGAGCATTT 180
DB 121 TACTGGGGGGTCCCGAGAGCTCCGAGGATCAGGAGCTCATTTTTCAGATCAGAGCATTT 180
QY 181 CAGAGAGCATTTTACCTACACCTTACGCGCGGATGCTCAGTTTGGCTCCCGCTTCTCC 240
DB 181 CAGAGAGCATTTTACCTACACCTTACGCGCGGATGCTCAGTTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCGCTCCAGAGGAGCTACGCGGGGCTCTTACAGACTGCGACGC 300
DB 241 ACTGAGCATCTGGGCGTCCCGCTCCAGAGGAGCTACGCGGGGCTCTTACAGACTGCGACGC 300
```

```

QY 301 TCGTCTATCTGGGAGCTGAGACCGGAGCGGACTGCTGCTGCTGAGCCGTGTC 360
DB 301 TCGTCTATCTGGGAGCTGAGACCGGAGCGGAGCGGACTGCTGCTGCTGAGCCGTGTC 360
QY 361 GGGGGGCTCCGGGAGCCTTTGGCTACCGAGGCGCCAGTATGTCATTAGCCCTGCCC 420
DB 361 GGGGGGCTCCGGGAGCCTTTGGCTACCGAGGCGCCAGTATGTCATTAGCCCTGCCC 420
QY 421 AATGCTACCGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 480
DB 421 AATGCTACCGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 480
QY 481 GGTGTTCCGGGCGGCGCTTCCGAGACCCCACTCTGCTCCGCGGAGTGGCCCTCGGCTG 540
DB 481 GGTGTTCCGGGCGGCGCTTCCGAGACCCCACTCTGCTCCGCGGAGTGGCCCTCGGCTG 540
QY 541 AACCCTCCATCTTACGCGGCGCTTACACCTTACAGCCGCGGCGGCGGCGGCTTGGGCGAG 600
DB 541 AACCCTCCATCTTACGCGGCGCTTACAGCCGCGGCGGCGGCGGCGGCGGCTTGGGCGAG 600
QY 601 AGTGTACCGCGGCGGAGGCTGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 660
DB 601 AGTGTACCGCGGCGGAGGCTGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 660
QY 661 GAGACGCTGTGTGTCGCGGAGAGTCAATGTCAGTTCAGAGCGGCGGAGCGGAGCGGAG 720
DB 661 GAGACGCTGTGTGTCGCGGAGAGTCAATGTCAGTTCAGAGCGGCGGAGCGGAGCGGAG 720
QY 721 TATCTGTGAGCGCTGCTGCGAAGCGGCGGAGCTTACCGCATCCAGATCTCTAAC 780
DB 721 TATCTGTGAGCGCTGCTGCGAAGCGGCGGAGCTTACCGCATCCAGATCTCTAAC 780
QY 781 CCCATCAACATGCTGTGTGTCAGAGTGTCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840
DB 781 CCCATCAACATGCTGTGTGTCAGAGTGTCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840
QY 841 GTCACCGGCAATGCGGCGCTGCGAAGCGGCGGAGCTTACCGCATCCAGATCTCTAAC 900
DB 841 GTCACCGGCAATGCGGCGCTGCGAAGCGGCGGAGCTTACCGCATCCAGATCTCTAAC 900
QY 901 AAGTGAAGTGAACACCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 960
DB 901 AAGTGAAGTGAACACCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 960
QY 961 CTGTGTGAGCGGAGCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1020
DB 961 CTGTGTGAGCGGAGCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1020
QY 1021 CCCAAGAGAGTGTGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1080
DB 1021 CCCAAGAGAGTGTGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1080
QY 1081 CAGGAGCTGGGCGGAGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1140
DB 1081 CAGGAGCTGGGCGGAGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1140
QY 1141 TTTGGGAACCTCCGAGCAACCAATGATGTCCCGGAGCTTATCCAGATGACCGTCCC 1200
DB 1141 TTTGGGAACCTCCGAGCAACCAATGATGTCCCGGAGCTTATCCAGATGACCGTCCC 1200
QY 1201 AACCCCTGTGTGAGCTGTGAGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1260
DB 1201 AACCCCTGTGTGAGCTGTGAGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1260
QY 1261 GACTGCTGTGTGAGCAACCAACCAATGATGTCCCGGAGCTTATCCAGATGACCGTCCC 1320
DB 1261 GACTGCTGTGTGAGCAACCAACCAATGATGTCCCGGAGCTTATCCAGATGACCGTCCC 1320
QY 1321 AGCTACACCTGAGCAGCAGTGTGAGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1380
DB 1321 AGCTACACCTGAGCAGCAGTGTGAGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1380
QY 1381 TACATGCACTACTGCAACCAAGTGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1440
DB 1381 TACATGCACTACTGCAACCAAGTGTGTGTCATTTGAGAGCATGGGCTTCCATCAGCCTTACCACTGCCC 1440
```



```
Db 1381 TACATGAGATGACACCAAGCTGTGTGTCACGGGAAAGCCAAAGGACATGTGTGC 1440
QY 1441 CAGACCCGACCTTCCCTGGCCGATGTGACAGCGTGGGAGGCAAGCTCTCC 1500
Db 1441 CAGACCCGACCTTCCCTGGCCGATGTGACAGCGTGGGAGGCAAGCTCTCC 1500
QY 1501 AAAGGGGCTGTGGAGAGACAACTTCAACAGCAGGCTGTATGCTTCTGAGCC 1560
Db 1501 AAAGGGGCTGTGGAGAGACAACTTCAACAGCAGGCTGTATGCTTCTGAGCC 1560
QY 1561 AAATGGGATCCCTATGGCCCTCTGCGGACATGTGTGGGGCTGTGACGCTGCGCAG 1620
Db 1561 AAATGGGATCCCTATGGCCCTCTGCGGACATGTGTGGGGCTGTGACGCTGCGCAG 1620
QY 1621 AGGCAATGACCAACCCCAACCCCTGCGCAAGAGTGTGAGGAGAGGAGG 1680
Db 1621 AGGCAATGACCAACCCCAACCCCTGCGCAAGAGTGTGAGGAGAGGAGG 1680
QY 1681 AAATACCGATCTGCAATGTGAGCCCTGCGCAGCTCAGCTCCGAAAGAGCTTCCG 1740
Db 1681 AAATACCGATCTGCAATGTGAGCCCTGCGCAGCTCAGCTCCGAAAGAGCTTCCG 1740
QY 1741 GAGGACAGTGTGAGCTTCAACGGCTCAACAGCAGCAACCGGCTCACTCTCGCC 1800
Db 1741 GAGGACAGTGTGAGCTTCAACGGCTCAACAGCAGCAACCGGCTCACTCTCGCC 1800
QY 1801 GTGGCATGGGTGCCAAGTACTCCGGCTGTCTCCCGGAGCAAGTCAAGCTCATCTGC 1860
Db 1801 GTGGCATGGGTGCCAAGTACTCCGGCTGTCTCCCGGAGCAAGTCAAGCTCATCTGC 1860
QY 1861 CGAGGCAATGGGACCTGCTCTCTATGTGTGTGACACCAAGGTGTGAGCGGCAAGCTG 1920
Db 1861 CGAGGCAATGGGACCTGCTCTCTATGTGTGTGACACCAAGGTGTGAGCGGCAAGCTG 1920
QY 1921 TGTCTCTGACTCTCACTCCCTCTGTGTGTGCAAGGCAAGTGTGAGGCTGGTGTAT 1980
Db 1921 TGTCTCTGACTCTCACTCCCTCTGTGTGTGCAAGGCAAGTGTGAGGCTGGTGTAT 1980
QY 1981 GGGAACTGGGCTCCCAAGAGAGATTCGACAGTGTGGGGTGTGTGGGGAGACAATAG 2040
Db 1981 GGGAACTGGGCTCCCAAGAGAGATTCGACAGTGTGGGGTGTGTGGGGAGACAATAG 2040
QY 2041 ACCTGCAAGAAGGTGACTGCTGCTTCAACAGCCATGCTATCAATTTCTGTGTG 2100
Db 2041 ACCTGCAAGAAGGTGACTGCTGCTTCAACAGCCATGCTATCAATTTCTGTGTG 2100
QY 2101 GCCATCCCGGAGGCGCTCAAGATCGACATCCGCAAGCGGTGTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGGAGGCGCTCAAGATCGACATCCGCAAGCGGTGTACAAAGGCTGATC 2160
QY 2161 GGGGATGACACTACTGCTGTGAAGAAACAGCAAGGCAAGTACTCTCAACGGGCAAT 2220
Db 2161 GGGGATGACACTACTGCTGTGAAGAAACAGCAAGGCAAGTACTCTCAACGGGCAAT 2220
QY 2221 TTTCGTGTGTGGGGTGGAGGCGGAGCGTGTGAAGGAGTGTGTGTGTGTGTGTGTGT 2280
Db 2221 TTTCGTGTGTGGGGTGGAGGCGGAGCGTGTGAAGGAGTGTGTGTGTGTGTGTGTGT 2280
QY 2281 GGCACGGGACAGCGGTGAGAGCTGTGAGAGCTTCCGCGCCATCTGTGACCGGTGACC 2340
Db 2281 GGCACGGGACAGCGGTGAGAGCTGTGAGAGCTTCCGCGCCATCTGTGACCGGTGACC 2340
QY 2341 GTGGAGGTCTCTCTGCTGGGAGAGTGAACACCGCCGGGTCTCTCTCTCTCTCTCTCT 2400
Db 2341 GTGGAGGTCTCTCTGCTGGGAGAGTGAACACCGCCGGGTCTCTCTCTCTCTCTCTCTCT 2400
QY 2401 CCAAGAGAGCTCTGGGAGAGAGTCTCTCATCCCAAGGAGCCCGGGGAGCCCTCTGTCT 2460
Db 2401 CCAAGAGAGCTCTGGGAGAGAGTCTCTCATCCCAAGGAGCCCGGGGAGCCCTCTGTCT 2460
QY 2461 TTGCAACAAGCGTCTCAAGCTCTCAACAGGTGAGAGCGGAGAGAGAGCCCT 2520
Db 2461 TTGCAACAAGCGTCTCAAGCTCTCAACAGGTGAGAGCGGAGAGAGAGAGCCCT 2520
```

```
Db 2461 TTGCAACAAGCGTCTCAAGCTCTCAACAGGTGAGAGCGGAGAGAGAGCCCT 2520
QY 2521 GCACGCTGGGTGGTGGAGAGCTGGGGGCGGTGTCTCCGAGCTGGGAGTGGCTGAG 2580
Db 2521 GCACGCTGGGTGGTGGAGAGCTGGGGGCGGTGTCTCCGAGCTGGGAGTGGCTGAG 2580
QY 2581 AAGCGGGGCTGTGAGTGTGCGGGGCTCCGCGGGGAGCGGAGCGTGTCTGTGTGTGT 2640
Db 2581 AAGCGGGGCTGTGAGTGTGCGGGGCTCCGCGGGGAGCGGAGCGTGTGTGTGTGTGT 2640
QY 2641 GCCATCGGCGCGGTGGAGACAAAGCTGTGCGGGAGAGCCCTCCGAGCTGGAGCTGAGC 2700
Db 2641 GCCATCGGCGCGGTGGAGACAAAGCTGTGCGGGAGAGCCCTCCGAGCTGGAGCTGAGC 2700
QY 2701 GCTGTGTACCTGTCTCAAGAGCTGTGCGGGGAGATTTCAAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCTGTGTACCTGTCTCAAGAGCTGTGCGGGGAGATTTCAAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCAAGGAGCGGCTGTGAGCGGGGAGCGGAGCTGCACTTGTGACCGGAGCG 2820
Db 2761 GTGGGCAAGGAGCGGCTGTGAGCGGGGAGCGGAGCTGCACTTGTGACCGGAGCG 2820
QY 2821 GAGCTGACCTTCTGCTCTGAGGCGGCTGTA 2853
Db 2821 GAGCTGACCTTCTGCTCTGAGGCGGCTGTA 2853
```

RESULT 2

US-09-965-631-7

Sequence 7, Application US/0995631

GENERAL INFORMATION:

APPLICANT: Friddele, Carl Johan

APPLICANT: Hilbun, Erin

TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same

FILE REFERENCE: LEX-0241-USA

CURRENT FILING DATE: 2001-09-27

PRIOR APPLICATION NUMBER: US 60/236,689

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 3446

TYPE: DNA

ORGANISM: homo sapiens

US-09-965-631-7

Query Match 100.0%; Score 2853; DB 36; Length 3446;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCTTCTGCTGGGATCTTAACCTGCTTGTGCGCGGGAACCGTGTGAGGCTGTGAG 60
Db 397 ATGCTTCTGCTGGGATCTTAACCTGCTTGTGCGCGGGAACCGTGTGAGGCTGTGAG 456
QY 61 CCAGAGGGGAGGTAGTGTGCTTCCATCCGACTGACCCGAGACATTAACGCGCCGCTAC 120
Db 457 CCAGAGGGGAGGTAGTGTGCTTCCATCCGACTGACCCGAGACATTAACGCGCCGCTAC 516
QY 121 TACTGGGGGGTCCCGAGAGTCCGGGAGTCAAGGACTATTTTCAGATCAGAGCATTT 180
Db 517 TACTGGGGGGTCCCGAGAGTCCGGGAGTCAAGGACTATTTTCAGATCAGAGCATTT 576
QY 181 CAGAGGACTTTTACCTACCTGACGCGGAGTGTGCTTGTGCTCCCGCTTCTCC 240
Db 577 CAGAGGACTTTTACCTACCTGACGCGGAGTGTGCTTGTGCTCCCGCTTCTCC 636
QY 241 ACTGAGCATGTGGGCGTCCCGCTCAAGGGGCTCAAGGGGCGCTTCAAGAGCTGCGAGCG 300
Db 637 ACTGAGCATGTGGGCGTCCCGCTCAAGGGGCTCAAGGGGCGCTTCAAGAGCTGCGAGCG 696
QY 301 TGTCTTATTTCTGGGAGCGTGAACGCGGAGCGGAGCTGCTGCTGTGAGCTGTGC 360
Db 301 TGTCTTATTTCTGGGAGCGTGAACGCGGAGCGGAGCTGCTGCTGTGAGCTGTGC 360
```

Db 697 TCGTTTATTTCTGGGAGCTGAACCGCCGAGCTGTTGGCTGCTGTGACCTGTGC 756
Oy 361 GGGGGGCTCCGGGAGACCTTTTGGCTACCGAGCGCCGAGTATGTATTAGCCCGCTGCC 420
Db 757 GGGGGGCTCCGGGAGACCTTTTGGCTACCGAGCGCCGAGTATGTATTAGCCCGCTGCC 816
Oy 421 AATGCTAGGCGCGCGCGCGAGCGGCAACAGCCAGGGGCGACACTTCTCCAGGCGGG 480
Db 817 AATGCTAGGCGCGCGCGCGAGCGGCAACAGCGGCGACACTTCTCCAGGCGGG 876
Oy 481 GGTGTTCCGGGCGGGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCTCTGGGCTGG 540
Db 877 GGTGTTCCGGGCGGGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCTCTGGGCTGG 936
Oy 541 AACCCCGCATCTCTACGGGCGCTTGACCTTACAAACCGCGCGCGGGCTTTGGGGAG 600
Db 937 AACCCCGCATCTCTACGGGCGCTTGACCTTACAAACCGCGCGGGCTTTGGGGAG 996
Oy 601 AGTGTAGCGCGGCGAGGCTGAGGCGCGCAACGCTTGTGTCTATCCGCGGTACTG 660
Db 997 AGTGTAGCGCGGCGAGGCTTGGGCGCGCAACGCTTGTGTCTATCCCGGTACTG 1056
Oy 661 GAGACGCTGTGTGTCGCGGAGAGTCAATGTCAAGTTCACAGGCGCGGACCTGACAT 720
Db 1057 GAGACGCTGTGTGTCGCGGAGAGTCAATGTCAAGTTCACAGGCGCGGACCTGACAT 1116
Oy 721 TATCTGTACGCTGCTGCGCAACGGGCGGCGACTACCGGCATCCGACATCCCTCAAC 780
Db 1117 TATCTGTACGCTGCTGCGCAACGGGCGGCGACTACCGGCATCCGACATCCCTCAAC 1176
Oy 781 CCCATCAACATGTTGTGTCAAGGTGCTGCTCTTAGAGATGTAAGTCTGCGGCGCAAG 840
Db 1177 CCCATCAACATGTTGTGTCAAGGTGCTGCTCTTAGAGATGTAAGTCTGCGGCGCAAG 1236
Oy 841 GTCAACGCGCAATGCGGCGCTTGACGCTGCGCAACTTCTGTGCTGCGGAGAGAGCTGAC 900
Db 1237 GTCAACGCGCAATGCGGCGCTTGACGCTGCGCAACTTCTGTGCTGCGGAGAGAGCTGAC 1296
Oy 901 AAGGTAGTACAAAGCAACCCCGAGTACTGGGACACTGCACTCTCTTCAACAGGCGAGAC 960
Db 1297 AAGGTAGTACAAAGCAACCCCGAGTACTGGGACACTGCACTCTCTTCAACAGGCGAGAC 1356
Oy 961 CTGTGTGAGCCACACCTGTGACACCTTGGGCAATGGCTGATGGTACATGTGTGAC 1020
Db 1357 CTGTGTGAGCCACACCTGTGACACCTTGGGCAATGGCTGATGGTACATGTGTGAC 1416
Oy 1021 CCCAAGAGAGCTGCTGTGATTTGAGAGATGGGCTTCATCAGGCTTCAACACCTGCC 1080
Db 1417 CCCAAGAGAGCTGCTGTGATTTGAGAGATGGGCTTCATCAGGCTTCAACACCTGCC 1476
Oy 1081 CACGAGCTGGGCGCACTGTTCACATGCCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1140
Db 1477 CACGAGCTGGGCGCACTGTTCACATGCCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1536
Oy 1141 TTTGGGAAAGCTCGAGCCAAACACATATGTCCCCGACCTCATCCAGATGACGCTGCC 1200
Db 1537 TTTGGGAAAGCTCGAGCCAAACACATATGTCCCCGACCTCATCCAGATGACGCTGCC 1596
Oy 1201 AACCCGTGTCAGCTGAGTGAAGTGGCATCATCAGACTTCCGAGACCGGAGAGGT 1260
Db 1597 AACCCGTGTCAGCTGAGTGAAGTGGCATCATCAGACTTCCGAGACCGGAGAGGT 1556
Oy 1261 GACTGCTCTCTGACCAACCCAGCAAGCCCATCTCCCTGCCGAGATCTGCGGGGCGCC 1320
Db 1657 GACTGCTCTCTGACCAACCCAGCAAGCCCATCTCCCTGCCGAGATCTGCGGGGCGCC 1716
Oy 1321 AGCTACACCTTGAGCCAGCAAGTGGAGTGGCTTTTGGCGTGGCTCAAAGCCTGTCT 1380
Db 1717 AGCTACACCTTGAGCCAGCAAGTGGAGTGGCTTTTGGCGTGGCTCAAAGCCTGTCT 1776
Oy 1381 TACATGAGTACGTCACCAACGCTGTGGTGCACGGGAAGGCGAAGGAGACGATGTGTC 1440
Db 1777 TACATGAGTACGTCACCAACGCTGTGGTGCACGGGAAGGCGAAGGAGACGATGTGTC 1836

Oy 1441 CAGACCCCGCATCTCCCTGGGGCGATGGACCAAGCTGTGGCGAGGGCAAGCTCTGCTC 1500
Db 1837 CAGACCCCGCATCTCCCTGGGGCGATGGACCAAGCTGTGGCGAGGGCAAGCTCTGCTC 1896
Oy 1501 AAGGGGCTGCTGAGAGAGACACAACTCAACAGCACAGGGTGGATGGTCTTCTGGGCC 1560
Db 1897 AAGGGGCTGCTGAGAGAGACACAACTCAACAGCACAGGGTGGATGGTCTTCTGGGCC 1956
Oy 1561 AATGGGATCCCTATGCGCCCTGCTGCGCAATGTTGGGGGCGGCTGAGCTGGCCAGG 1620
Db 1957 AATGGGATCCCTATGCGCCCTGCTGCGCAATGTTGGGGGCGGCTGAGCTGGCCAGG 2016
Oy 1621 AGCGAGTACCAACCCACCCCTGCGCAAGGGGGGCAAGTACTGCGAGGAGTGGAGGTG 1680
Db 2017 AGCGAGTACCAACCCACCCCTGCGCAAGGGGGGCAAGTACTGCGAGGAGTGGAGGTG 2076
Oy 1681 AATACCGATCTCTCAATCTGAGGCGCTTCAAGGGGTACAAACAGCAACCGGCTCACTGCC 1740
Db 2077 AATACCGATCTCTCAATCTGAGGCGCTTCAAGGGGTACAAACAGCAACCGGCTCACTGCC 2136
Oy 1741 GAGGAGCTGTGAGGCGCTTCAAGGGGTACAAACAGCAACCGGCTCACTGCC 1800
Db 2137 GAGGAGCTGTGAGGCGCTTCAAGGGGTACAAACAGCAACCGGCTCACTGCC 2196
Oy 1801 GTGGCATGGTGGCCCAAGTACTCGGGGTGTCTCCCGGGCAAGTGCACAGCTTCATCTGC 1860
Db 2197 GTGGCATGGTGGCCCAAGTACTCGGGGTGTCTCCCGGGCAAGTGCACAGCTTCATCTGC 2256
Oy 1861 CGAGCAATGCTGCTGCTACTTCTATGCTGCGACCCCAAGGGTGGAGGCGAGCTG 1920
Db 2257 CGAGCAATGCTGCTGCTACTTCTATGCTGCGACCCCAAGGGTGGAGGCGAGCTG 2316
Oy 1921 TGTCTCTGACTCCACTCCGCTGCTGTGTCCCAAGGCAAGTGCATCAAGGCTGCTGAT 1980
Db 2317 TGTCTCTGACTCCACTCCGCTGCTGTGTCCCAAGGCAAGTGCATCAAGGCTGCTGAT 2376
Oy 1981 GGGAACTGGGCTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGCAATAG 2040
Db 2377 GGGAACTGGGCTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGCAATAG 2436
Oy 2041 AGCTGCAAGAGGCTGCTGCTGCTTCAACAGCCCATGCTATGCTCAATTTCTGGTGTG 2100
Db 2437 AGCTGCAAGAGGCTGCTGCTGCTTCAACAGCCCATGCTATGCTCAATTTCTGGTGTG 2496
Oy 2101 GCCATCCCGCAGGCGCTCAAGCATGCAATCCGCGAGCCGGGTACAAAGGCTGATC 2160
Db 2497 GCCATCCCGCAGGCGCTCAAGCATGCAATCCGCGAGCCGGGTACAAAGGCTGATC 2556
Oy 2161 GGGGATGACACTACTGCTGCTGTAAGAAACGCAAGGCAAGTACCTGTCACAGGGCAT 2220
Db 2557 GGGGATGACACTACTGCTGCTGTAAGAAACGCAAGGCAAGTACCTGTCACAGGGCAT 2616
Oy 2221 TTTGTTGTGTGGGCGGTGAGGCGGAGCTGTGTGAAGGGCAGTCTGCTCGGTACAGC 2280
Db 2617 TTTGTTGTGTGGGCGGTGAGGCGGAGCTGTGTGAAGGGCAGTCTGCTCGGTACAGC 2676
Oy 2281 GGCAGGGGACAGCGGTGAGAGCTGCAAGGCTTCCGAGCCCATCTTGGAGCCGTGACC 2340
Db 2677 GGCAGGGGACAGCGGTGAGAGCTGCAAGGCTTCCGAGCCCATCTTGGAGCCGTGACC 2736
Oy 2341 GTGGAGTCTCTGCTGGGGAAAGTACACCGCGCGGGTCCGCTACTCTCTATGTG 2400
Db 2737 GTGGAGTCTCTGCTGGGGAAAGTACACCGCGCGGGTCCGCTACTCTCTATGTG 2796
Oy 2401 CCCAAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGAGACCCCGGGGAGCCTCTGTC 2460
Db 2797 CCCAAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGAGACCCCGGGGAGCCTCTGTC 2856
Oy 2461 TTGCACACAGGCTCTGAGCTTCAACACAGTGTGAGAGCGGAGACAGAGCCCTCT 2520
Db 2857 TTGCACACAGGCTCTGAGCTTCAACACAGTGTGAGAGCGGAGACAGAGCCCTCT 2916

QY	2521	GCACGTGGGAGCGCTGGCACTCTGGGGGCGCGTCCGGCAAGCTGGGGCAGTGGCTGCAG	2580
Db	2917	GCACGCTGGGTGGCTGGCACTCTGGGGGCGCGTCTCGGCAAGCTGGGGCAGTGGCTGCAG	2978
QY	2581	AAGCGGGCGGTGACCTGTGGGGCTCCGGCGGGCAGCCGACCGATGCTCCCTGGTGTGATCA	2640
Db	2977	AAGGGGCGGTGACCTGTGGGGCTCCGGCGGGCAGCCGACCGATGCTCCCTGGTGTGATCA	3036
QY	2641	GCCCATGGGCCCTGGAGACACAAAGCTCTGGGGGAGCCCTGGCCGACCTGGGAGTCCG	2700
Db	3037	GCCCATGGGCCGTGGAGACACAAAGCTCTGGGGGAGCCCTGGCCGACCTGGGAGTCTAGC	3096
QY	2701	GCTCGGTACCCCTGCTCCAAAGACTCTGGCCCGGGGATTTCAAGAGCGCTCACTCAAGGT	2760
Db	3097	GCTCGGTACCCCTGCTCCAAAGACTCTGGCCCGGGGATTTCAAGAGCGCTCACTCAAGGT	3156
QY	2761	GTGGGCGACGAGAGCGCGCTGTGTGGCCCGGAGACAGTCAATTGACACGCAAGGCCAG	2820
Db	3157	GTGGGCGACGAGAGCGCGCGCTGTGTGGCCCGGAGACAGTCAATTGACACGCAAGGCCAG	3216
QY	2821	GAGTGGACTTGTGGCTCCCTGAGAGCGGCTCTCA	2853
Db	3217	GAGTGGACTTGTGGCTCCCTGAGAGCGGCTCTCA	3249

```

RESULT 3
US-09-741-151-1
: Sequence 1, Application US/09741151
: GENERAL INFORMATION:
: APPLICANT: ZHU, Shaoqing et al
: TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
: TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: CL001005
: CURRENT APPLICATION NUMBER: US/09/741,151
: CURRENT FILING DATE: 2000-12-08
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 2867
: TYPE: DNA
: ORGANISM: Human
: US-09-741-151-1

```

Query Match	94.68;	Score 2700;	DB 29;	Length 2867;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 2850; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

OY	1	ATGTTTGTGGGGACCTTAACCCGGGCTTTCGCCGGGCGAAACCGCTGAGAGCTCTGAG	60
Db	4	ATGCTTGTGTGGGCACTTAACCCCTGGCTTTCGCCGGGCGAAACCCGCTGAGAGCTCTGAG	63
OY	61	CGAGAGCGGAGAGTATGCTTTCCTTCGAGCTGGACCCGGACATTTAAGCGCCGCGCTAC	120
Db	64	CGAGAGCGGAGAGTATGCTTTCCTTCGAGCTGGACCCGGACATTTAAGCGCCGCGCTAC	122
OY	121	TACTGGCGGGGTCGCCGAGACTCCGGGGATCAGGACTATTTTTCAGATCAGACATTT	180
Db	124	TACTGGCGGGGTCGCCGAGACTCCGGGGATCAGGACTATTTTTCAGATCAGACATTT	183
OY	181	CAGSAGGACTTTTACTCTACCTACGTAGCGCGGAGTGTACGTTCTTGAGCTCCGCGCTTCC	240
Db	184	CAGSAGGACTTTTACTCTACCTACGTAGCGCGGAGTGTACGTTCTTGAGCTCCGCGCTTCC	243
OY	241	ACTGAGCATCTTGGGGTCCCGCTCCAGGGGCTACCGGGGGGCTTTCAAGCTGGGAAGC	300
Db	244	ACTGAGCATCTTGGGGTCCCGCTCCAGGGGCTACCGGGGGGCTTTCAAGCTGGGAAGC	303
OY	301	TGCTTTATTTCTGGGGAGGTGAACGCCGAGACGAGTCTGTTGCTGCTGTGAGCTGTGC	360
Db	304	TGCTTTATTTCTGGGGAGGTGAACGCCGAGACGAGTCTGTTGCTGCTGTGAGCTGTGC	363
OY	361	GGGGGGGCTCCGGGAGCCTTTGGCTTACCGAGGCGCGAGTATGTATTAAGCCCGCTGCC	420

Db	364	GGGGGGCTCGGGAGACCTTTTGGCTACCGAGAGCGCGAGTGTTCATTAGCCCAATGGCC	423
Qy	421	AAATGCTAGGGCGCGCGCGCGCGAGCGCAACAGCCAGGGCGGCACACCTTCTCCACCGCCG	480
Db	424	AAATGCTAGGGCGCGCGCGCGCGAGCGCAACAGCCAGGGCGGCACACCTTCTCCACCGCCG	483
Qy	481	GGTGTTCGGGGGGGGGCTTTCGGAGAACCCCACTCTGCTGCGGGGGTGGCGCTGGGGCTGG	540
Db	484	GGTGTTCGGGGGGGGGCTTTCGGAGAACCCCACTCTGCTGCGGGGGTGGCGCTGGGGCTGG	543
Qy	541	AAACCCGCCATCTCTACGGGCGCTTGAGACCTTTACAAAGCGCGGGCGGGCGCTTC	600
Db	544	AAACCCGCCATCTCTACGGGCGCTTGAGACCTTTACAAAGCGCGGGCGGGCGCTTCGGGGAG	603
Qy	601	AGTCGAGAGCGGCGAGGTCGTGGGGGGCGCCCAAGGGTTTCGTCTATACCGGGGTACGTG	660
Db	604	AGTCGAGAGCGGCGAGGTCGTGGGGGGCGCCCAAGGGTTTCGTCTATACCGGGGTACGTG	663
Qy	661	GAGAGCGTGTGTGCGCGAGAGTCAATGTGTCAAGTTCCACGGCGCGGACCTGTGAACAT	720
Db	664	GAGAGCGTGTGTGCGCGAGAGTCAATGTGTCAAGTTCCACGGCGCGGACCTGTGAACAT	723
Qy	721	TATCTGCTGAGGCTGCTGGCAAGCGGCGCGCGAGCTTACCGGCCATCCAGACATCTCAAC	780
Db	724	TATCTGCTGAGGCTGCTGGCAAGCGGCGCGAGCTTACCGGCCATCCAGACATCTCAAC	783
Qy	781	CCCATCAACATGTTGTGTGAAGTGTGCTTTTGAAGATGATGTGACTCCGGGCCCAAG	840
Db	784	CCCATCAACATGTTGTGTGAAGTGTGCTTTTGAAGATGATGTGACTCCGGGCCCAAG	843
Qy	841	GTACACGGCAATGCGGGCGCTTACGCTGGCACTTCTGTGCTGGCAGAAAGATGCAAC	900
Db	844	GTACACGGCAATGCGGGCGCTTACGCTGGCACTTCTGTGCTGGCAGAAAGATGCAAC	903
Qy	901	AAAGTAGTAGAACACACCCCGAGACTGGGGACACTGCGCATCTCTTCCACAGGACGAC	960
Db	904	AAAGTAGTAGAACACACCCCGAGACTGGGGACACTGCGCATCTCTTCCACAGGACGAC	963
Qy	961	CTGTGTGAGCCACCACTGTGACACCCCTGGGCAATGGCTATGTGGTACCATGTGTGAC	1020
Db	964	CTGTGTGAGCCACCACTGTGACACCCCTGGGCAATGGCTATGTGGTACCATGTGTGAC	1023
Qy	1021	CCCAAGAGAAGCTGTCTGTATTTAGSAGATGGGCTTCCATCAAGCCTTTCACCACTGCGC	1080
Db	1024	CCCAAGAGAAGCTGTCTGTATTTAGSAGATGGGCTTCCATCAAGCCTTTCACCACTGCGC	1083
Qy	1081	CACGAGCTGGGGCCAGTGTTCATCATCCTCCCATGACAAATGTGAATCTGTGACGAGTG	1140
Db	1084	CACGAGCTGGGGCCAGTGTTCATCATCCTCCCATGACAAATGTGAATCTGTGACGAGTG	1143
Qy	1141	TTTGGGAAAGCTCCAGCCACACACATGATGTCCCGAACCCCTCATCCAGATGACAGCTGCC	1200
Db	1144	TTTGGGAAAGCTCCAGCCACACACATGATGTCCCGAACCCCTCATCCAGATGACAGCTGCC	1203
Qy	1201	AAACCCCTGTAGGCTGTGAGTGGCTGCATATACGAGACTTCCTGTGACGCGGGACAGGT	1260
Db	1204	AAACCCCTGTAGGCTGTGAGTGGCTGCATATACGAGACTTCCTGTGACGCGGGACAGGT	1263
Qy	1261	GACTGCTCTTGGACCAACCCAGACACCCATCTCCCTGCCCGAGAGATTTGCTGCGCGCC	1320
Db	1264	GACTGCTCTTGGACCAACCCAGACACCCATCTCCCTGCCCGAGAGATTTGCTGCGCGCC	1323
Qy	1321	AGCTACACCCGTGAGGCGAGGTGCGAGTGGGCTTTTGGGGTGGGCTCCAAAGCCCTGTCTCT	1380
Db	1324	AGCTACACCCGTGAGGCGAGGTGCGAGTGGGCTTTTGGGGTGGGCTCCAAAGCCCTGTCTCT	1383
Qy	1381	TACATGCAAGTACTGCAACAAGCTGTGTGTCACCGGGAAGGCAAGGACAGATGGTGTGC	1440
Db	1384	TACATGCAAGTACTGCAACAAGCTGTGTGTCACCGGGAAGGCAAGGACAGATGGTGTGC	1443
Qy	1441	CAGACCGCCACTTCCCCTGGGGCCGATGGCACACAGCTGTGGCGAGGGCAAGCTCTGCCCTC	1500

Db 1444 CAGACCCGACCTTCCCTGGGCGATGGACACGCTGT3CGAGAGGCAAGCTTGCCCTC 1503
 QY 1501 AAGGGGCGCTGGCTGGAGAGACAACTTCAACACAGAGGTGATGTTCTTGCGGC 1560
 Db 1504 AAGGGGCGCTGGCTGGAGAGACAACTTCAACACAGAGGTGATGTTCTTGCGGC 1563
 QY 1561 AATGGGATCCCTATGAGGCGCTGCGGCACATGTTGGGGGCTGACAGTGGCCAG 1620
 Db 1564 AATGGGATCCCTATGAGGCGCTGCGGCACATGTTGGGGGCTGACAGTGGCCAG 1623
 QY 1621 AGGAGTGACCAACCCACCCCTGCCAAGGGGGCAAGTACTGGAGGAGTGAAGGTG 1680
 Db 1624 AGGAGTGACCAACCCACCCCTGCCAAGGGGGCAAGTACTGGAGGAGTGAAGGTG 1683
 QY 1681 AATATCCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCCTCCGAAAGAGCTCCGG 1740
 Db 1684 AATATCCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCCTCCGAAAGAGCTCCGG 1743
 QY 1741 GAGGAGCATGTGAGGCTTTCAACGGGCTACACACAGCAGCAACCGGCTCACTCTCGCC 1800
 Db 1744 GAGGAGCATGTGAGGCTTTCAACGGGCTACACACAGCAGCAACCGGCTCACTCTCGCC 1803
 QY 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTGCACAGCTCATCTGC 1860
 Db 1804 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTGCACAGCTCATCTGC 1863
 QY 1861 CGAGCCATGAGCACTGCTACTTCTATGTGCTGGACCCCAAGGTGGTGGAGCGCAGCTG 1920
 Db 1864 CGAGCCATGAGCACTGCTACTTCTATGTGCTGGACCCCAAGGTGGTGGAGCGCAGCTG 1923
 QY 1921 TGCTCTCTGACTGACCTCCGCTGTGTGTCGAAGGCAAGTGCATCAAGCTGCTGTGAT 1980
 Db 1924 TGCTCTCTGACTGACCTCCGCTGTGTGTCGAAGGCAAGTGCATCAAGCTGCTGTGAT 1983
 QY 1981 GGGAGCTGGGCTCCAGAGAGATTCGACAGTGGGGTGGTGGGGAGAGCAATGAG 2040
 Db 1984 GGGAGCTGGGCTCCAGAGAGATTCGACAGTGGGGTGGTGGGGAGAGCAATGAG 2043
 QY 2041 AGCTGCAAGAAGTGTGACTGTGACTTTCACCAAGGCCATGATGATGATCAATTTGATG 2100
 Db 2044 AGCTGCAAGAAGTGTGACTGTGACTTTCACCAAGGCCATGATGATGATGATGATG 2103
 QY 2101 GCCATCCCCGAGCGGCTCAAGCATGACATCGCCAGCGGTTTACAAAGGCTGATC 2160
 Db 2104 GCCATCCCCGAGCGGCTCAAGCATGACATCGCCAGCGGTTTACAAAGGCTGATC 2163
 QY 2161 GGGATGACAACTACTGCGCTCTGAAGACAGCCAGCAAGTACCTGCTCAACGGGCA 2220
 Db 2164 GGGATGACAACTACTGCGCTCTGAAGACAGCCAGCAAGTACCTGCTCAACGGGCA 2223
 QY 2221 TTGCTGTGTGCGGCGGTGAGACGGGACCTGCTGTGTAAGGAGTCTGCTGCGGTACAGC 2280
 Db 2224 TTGCTGTGTGCGGCGGTGAGACGGGACCTGCTGTGTAAGGAGTCTGCTGCGGTACAGC 2283
 QY 2281 GGCACGGGCAAGGGGTGGAGAGCTTCAAGGCTTCCGGCCCAATCCGTAAGCGGTAGC 2340
 Db 2284 GGCACGGGCAAGGGGTGGAGAGCTTCAAGGCTTCCGGCCCAATCCGTAAGCGGTAGC 2343
 QY 2341 GTGAGGCTCTCTCCGTGGGGAATGACACCGCGGGGTCCGCTACTCTTCTATCTG 2400
 Db 2344 GTGAGGCTCTCTCCGTGGGGAATGACACCGCGGGGTCCGCTACTCTTCTATCTG 2403
 QY 2401 CCCAAAGAGCTCGGAGAGACAGTCTCTGATCCCAAGACCCCGGGGACCCCTCTGTC 2460
 Db 2404 CCCAAAGAGCTCGGAGAGACAGTCTCTGATCCCAAGACCCCGGGGACCCCTCTGTC 2463
 QY 2461 TTGCAACACAGCTCTCTAGCTCTCAACAGGTGGAGAGCGCGGAGACAGAGCCCT 2520
 Db 2464 TTGCAACACAGCTCTCTAGCTCTCAACAGGTGGAGAGCGCGGAGACAGAGCCCT 2523
 QY 2521 GCACGCTGGTGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCT 2580
 Db 2524 GCACGCTGGTGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCT 2583

QY 2581 AAGCGGGGCTGAGTGTGCGGGCTCCGCGGAGCGACAGGTGCTTGTGATGCA 2640
 Db 2584 AAGCGGGGCTGAGTGTGCGGGGCTCCGCGGAGCGACAGGTGCTTGTGATGCA 2643
 QY 2641 GCCCATCGGGCGGTGAGACACAAAGCTCGGGGAGAGCCCTTCCCTACTGGAGCTAGC 2700
 Db 2644 GCCCATCGGGCGGTGAGACACAAAGCTCGGGGAGAGCCCTTCCCTACTGGAGCTAGC 2703
 QY 2701 GCTGTGACCTGCTCCAAAGCTGCGGGCGGGGATTTGAGAGGCTCACTCAAGTGT 2760
 Db 2704 GCTGTGACCTGCTCCAAAGCTGCGGGCGGGGATTTGAGAGGCTCACTCAAGTGT 2763
 QY 2761 GTGGGCAAGAGAGCGGCTGCTGCGGGGACCAAGTGCACCTTGACCCGAAAGCC 2820
 Db 2764 GTGGGCAAGAGAGCGGCTGCTGCGGGGACCAAGTGCACCTTGACCCGAAAGCC 2823
 QY 2821 GAGCTGAGCTTCTGCTCTTCAAGCCGCTGCTGA 2853
 Db 2824 GAGCTGAGCTTCTGCTCTTCAAGCCGCTGCTGA 2856

RESULT 4

US-10-009-332-2

; Sequence 2, Application US/10009332
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
 ; APPLICANT: Kazusa DNA Research Institute
 ; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY
 ; FILE REFERENCE: 067541
 ; CURRENT APPLICATION NUMBER: US/10/009, 332
 ; CURRENT FILING DATE: 2001-12-10
 ; PRIOR APPLICATION NUMBER: JPA Hei 11-321740
 ; PRIOR FILING DATE: 1999-11-11
 ; PRIOR APPLICATION NUMBER: JPA 2000-144020
 ; PRIOR FILING DATE: 2000-05-16
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 2853
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-009-332-2

Query Match 94.5%; Score 2695; DB 38; Length 2853;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 57 TGAGCGAGAGCGGAGAGTGTGCTTCCCATCCGACTGGACCCGAGCATTAAGGCGCGC 116
 Db 57 TGAGCGAGAGCGGAGAGTGTGCTTCCCATCCGACTGGACCCGAGCATTAAGGCGCGC 116
 QY 117 CTACTACTGCGGGGTCCCGAGGACTCGGGGATGAGGACTATTTTCAGATACAGC 176
 Db 117 CTACTACTGCGGGGTCCCGAGGACTCGGGGATGAGGACTATTTTCAGATACAGC 176
 QY 117 ATTTCAGAGGACTTTTACTTACTTACACTGACGGCGGATGATGATGCTTGGCTCCGCC 236
 Db 117 ATTTCAGAGGACTTTTACTTACTTACACTGACGGCGGATGATGATGCTTGGCTCCGCC 236
 QY 237 CTCGACTGAGCACTGCGGGCTCCCTCCAGGGGCTCAACCGGGGCTTTAGACCTGCG 296
 Db 237 CTCGACTGAGCACTGCGGGCTCCCTCCAGGGGCTCAACCGGGGCTTTAGACCTGCG 296
 QY 297 AGCGTCTTCTATTTCTGAGGAGTGAACGCGAGCGGACTGTTGCTGCTGTGAGCTT 356
 Db 297 AGCGTCTTCTATTTCTGAGGAGTGAACGCGAGCGGACTGTTGCTGCTGTGAGCTT 356
 QY 357 GTGCGGGGGGCTCCGGAGAGCTTTGCTTACAGAGGCGCGATGATTAAGCCGCT 416
 Db 357 GTGCGGGGGGCTCCGGAGAGCTTTGCTTACAGAGGCGCGATGATTAAGCCGCT 416
 QY 417 GCCCAATGCTAGCGCGCGCGGCGAGCGCAACAGCCAGGGCGGACACCTTCTCCAGCG 476

Db 417 GCCCAATGCTAGGCGCGCGCGCGCAAGCCAAAGCCAGGCGGCAACCTCTCCAGCG 476
OY 477 CCGGGGCTTTCCGGGGGCGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCCCTCGG 536
Db 477 CCGGGGCTTTCCGGGGGCGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCCCTCGG 536
OY 537 CTGGAAACCCCGCATCTCTACGGGCGCTTGACCTTACAAAGCCGCGGCGGGGCTTCG 596
Db 537 CTGGAAACCCCGCATCTCTACGGGCGCTTGACCTTACAAAGCCGCGGCGGGGCTTCG 596
OY 597 GAGAGCTCGTAGCCGCGGAGGCTTGCGGCGCGCCAAAGCTTTCTATCCCGCGTA 656
Db 597 GAGAGCTCGTAGCCGCGGAGGCTTGCGGCGCGCCAAAGCTTTCTATCCCGCGTA 656
OY 657 CGGAGAGCTGTGGTTCGCGAGAGTCAATGATAGTTCACAGCGCGGCGGACTGGA 716
Db 657 CGGAGAGCTGTGGTTCGCGAGAGTCAATGATAGTTCACAGCGCGGCGGACTGGA 716
OY 717 ACATTAATCTGCTAGCCTGCTGCAACGGCGCGGCTTACCGCATCCAGCATCT 776
Db 717 ACATTAATCTGCTAGCCTGCTGCAACGGCGCGGCTTACCGCATCCAGCATCT 776
OY 777 CAACCCCATTAACATGCTTGTGTCAAGGTGCTTCTTAAAGATCTGACTCCGCGC 836
Db 777 CAACCCCATTAACATGCTTGTGTCAAGGTGCTTCTTAAAGATCTGACTCCGCGC 836
OY 837 CAAGGTACCGGCAATGCGGCGCTGACGCTGGCGCACTTCTGCGCTGCGCAAGAGCT 896
Db 837 CAAGGTACCGGCAATGCGGCGCTGACGCTGGCGCACTTCTGCGCTGCGCAAGAGCT 896
OY 897 GAACAAAGTAGTAGAAGAGACCCCGAGTACTGGAGACATGCCATCTTTCACAGGCA 956
Db 897 GAACAAAGTAGTAGAAGAGACCCCGAGTACTGGAGACATGCCATCTTTCACAGGCA 956
OY 957 GAGACCTGTGTGGAGCCACCATCTGTACACCTGGCGATGCTGATGCGGTACCATGTG 1016
Db 957 GAGACCTGTGTGGAGCCACCATCTGTACACCTGGCGATGCTGATGCGGTACCATGTG 1016
OY 1017 TGACCCCAAGAGAGTGTCTGTCTATGAGAGCATGGGCTTCCATCGACCTTCACCG 1076
Db 1017 TGACCCCAAGAGAGTGTCTGTCTATGAGAGCATGGGCTTCCATCGACCTTCACCG 1076
OY 1077 TGCCCAAGAGCTGGCGCACAGTGTTCACATGCCCATGACCAATGTGAAAGCTGTGAGA 1136
Db 1077 TGCCCAAGAGCTGGCGCACAGTGTTCACATGCCCATGACCAATGTGAAAGCTGTGAGA 1136
OY 1137 GGTGTTGGGAAGCTCCGAGCCAAACATGATGTCCCGACCTTCATCCAGATGAGCG 1196
Db 1137 GGTGTTGGGAAGCTCCGAGCCAAACATGATGTCCCGACCTTCATCCAGATGAGCG 1196
OY 1197 TGCCCAACCCCTGTGAGCTGTGAGTGTCCATCATCAGCACTTCTGGAGAGCGGCA 1256
Db 1197 TGCCCAACCCCTGTGAGCTGTGAGTGTCCATCATCAGCACTTCTGGAGAGCGGCA 1256
OY 1257 CGGTGACTGCTCTGTGAGCAACCCAGAGCCCATCTCCCTGCGGAGATCTCCCGG 1316
Db 1257 CGGTGACTGCTCTGTGAGCAACCCAGAGCCCATCTCCCTGCGGAGATCTCCCGG 1316
OY 1317 CGGCACCTACACCTGAGCGAGAGTGTGAGTGTGCGGTGCGGTTCGAAGCCCTG 1376
Db 1317 CGGCACCTACACCTGAGCGAGAGTGTGAGTGTGCGGTGCGGTTCGAAGCCCTG 1376
OY 1377 TCCTTACATGAGTACTGACCAAGTGTGTGACACCGGAGAGGCAAGGACAGATGAT 1436
Db 1377 TCCTTACATGAGTACTGACCAAGTGTGTGACACCGGAGAGGCAAGGAGCAATGAT 1436
OY 1437 GTGCCAGACCCGCACTTCCCTGGCGGATGACACAGCTGTGGCGAGGCGCAAGCTGTG 1496
Db 1437 GTGCCAGACCCGCACTTCCCTGGCGGATGACACAGCTGTGGCGAGGCGCAAGCTGTG 1496
OY 1497 CCTCAAGGGGCGCTGTGGAGAGACAACTCAACAAAGCAAGGCTGATGCTCTG 1556
Db 1497 CCTCAAGGGGCGCTGTGGAGAGACAACTCAACAAAGCAAGGCTGATGCTCTG 1556

Db 1497 CCTCAAGGGGCGCTGTGGAGAGACAACTCAACAAAGCAAGGCTGATGCTCTG 1556
OY 1557 GGCCAAAATGGATTCCTTATGGCCCTGCTCGCGCACATGTGTGGGGCGGTGACGTTGC 1616
Db 1557 GGCCAAAATGGATTCCTTATGGCCCTGCTCGCGCACATGTGTGGGGCGGTGACGTTGC 1616
OY 1617 CAGGAGGCGAGTGCACCAACCCGCTGCGCAACCGGGGCAAGTCTGCGAGGGGAGTA 1676
Db 1617 CAGGAGGCGAGTGCACCAACCCGCTGCGCAACCGGGGCAAGTCTGCGAGGGGAGTA 1676
OY 1677 GGTGAATAACGATCTCTCAATCTGAGACCCCTGCCAGCTCAGCTTCGGAAAGCTT 1736
Db 1677 GGTGAATAACGATCTCTCAATCTGAGACCCCTGCCAGCTCAGCTTCGGAAAGCTT 1736
OY 1737 CCGGAGAGGAGCATGTGTGAGGCTTTCACAGGCTACACACAGCACAACCCGCTACCT 1796
Db 1737 CCGGAGAGGAGCATGTGTGAGGCTTTCACAGGCTACACACAGCACAACCCGCTACCT 1796
OY 1797 CGCGTGGCATGGGCGCCCAAGTACTCCGGGCTGTCTCCGGGGCAAGTGCAGCTCAT 1856
Db 1797 CGCGTGGCATGGGCGCCCAAGTACTCCGGGCTGTCTCCGGGGCAAGTGCAGCTCAT 1856
OY 1857 CTGCGAGCCCAATGSCACTGGCTACTTCTATGTGCTGCGCACCAAGTGTGACGCGAC 1916
Db 1857 CTGCGAGCCCAATGSCACTGGCTACTTCTATGTGCTGCGCACCAAGTGTGACGCGAC 1916
OY 1917 GCTGTGCTCTCTGTGACTCCTCCTGCTGTGTGTCACAGGCAAGTGCATCAAGCTGCTG 1976
Db 1917 GCTGTGCTCTCTGTGACTCCTCCTGCTGTGTGTCACAGGCAAGTGCATCAAGCTGCTG 1976
OY 1977 TGATGGGAACCTGGGCTCTCAAGAAAGATGACAAAGTGTGGGCTGTGGGAGAGCAA 2036
Db 1977 TGATGGGAACCTGGGCTCTCAAGAAAGATGACAAAGTGTGGGCTGTGGGAGAGCAA 2036
OY 2037 TAAGAGCTGCAAGAGTGTGACTGACTCTTACCAGAGCCATGCATGACTCAATTTCT 2096
Db 2037 TAAGAGCTGCAAGAGTGTGACTGACTCTTACCAGAGCCATGCATGACTCAATTTCT 2096
OY 2097 GGTGGCATCCCGGAGGCGCTCAAGATGACATCCGCGGCGGTTCACAAAGGCT 2156
Db 2097 GGTGGCATCCCGGAGGCGCTCAAGATGACATCCGCGGCGGTTCACAAAGGCT 2156
OY 2157 GATCGGGATGACAACATCTGCTGTGTAAGAAAGCAAGCAAGCAAGTACCTGCTCAACG 2216
Db 2157 GATCGGGATGACAACATCTGCTGTGTAAGAAAGCAAGCAAGCAAGTACCTGCTCAACG 2216
OY 2217 GCATTTCGTGTGTGCGGGGTGAGACCGGACCTGTGTGTAAGAGGAGTGTGCGGTA 2276
Db 2217 GCATTTCGTGTGTGCGGGGTGAGACCGGACCTGTGTGTAAGAGGAGTGTGCGGTA 2276
OY 2277 CAGCGGACCGGAGAGGCTGTGAGAGCTTCAGAGCTTCCCGGCGCATCCGTAAGCGCT 2336
Db 2277 CAGCGGACCGGAGAGGCTGTGAGAGCTTCAGAGCTTCCCGGCGCATCCGTAAGCGCT 2336
OY 2337 GACCGTGAAGTCTCTCCGTTGGGAAAGTACACCGCGCGGCTGCCACTCTCTCTA 2396
Db 2337 GACCGTGAAGTCTCTCCGTTGGGAAAGTACACCGCGCGGCTGCCACTCTCTCTA 2396
OY 2397 TCTGCGCCAAAGAGCTGTGGAGAGCAAGTCTCTCATCTCCAAAGAACCCCGGGAGCTTC 2456
Db 2397 TCTGCGCCAAAGAGCTGTGGAGAGCAAGTCTCTCATCTCCAAAGAACCCCGGGAGCTTC 2456
OY 2457 TGTCTTGACAAAGCGTCTCAGCTCTCCCAACAGAGTGGAGAGCGGAGCAAGCGC 2516
Db 2457 TGTCTTGACAAAGCGTCTCAGCTCTCCCAACAGAGTGGAGAGCGGAGCAAGCGC 2516
OY 2517 CCTTGACAGCTGGTGTGCTGCACTGTGGGCGCTGTCTCCGAGAGTGCAGAGTGGCT 2576
Db 2517 CCTTGACAGCTGGTGTGCTGCACTGTGGGCGCTGTCTCCGAGAGTGCAGAGTGGCT 2576
OY 2577 GCAGAAGCGGCGGTGAGTCTCGGGGCTCCGCGGAGCAGCAGGTCCTCTGCTGTGA 2636
Db 2577 GCAGAAGCGGCGGTGAGTCTCGGGGCTCCGCGGAGCAGCAGGTCCTCTGCTGTGA 2636

OY	181	CATAGAGACTTTAACTACACACCGAGCGCGGAAATGGCTCAATCTTGGATCCCGCCCTCTCC	240
Db	255	CAGAGAGACTTTAACTACACACCGAGCGCGGAAATGGCTCAATCTTGGATCCCGCCCTCTCC	314
OY	241	ACTAGAGATCTGGGCGCGCCCTCCACAGGGGCTCACCGGGGGCTCTTCAGACCTGACGCG	300
Db	315	ACTAGAGATCTGGGCGCGCCCTCCACAGGGGCTCACCGGGGGCTCTTCAGACCTGACGCG	374
OY	301	TGCTTTATTTCTGGGGACGTGAACGCCGAGCCGAGACTCTTGGCTGCTGTAGCCTGTGC	360
Db	375	TGCTTTATTTCTGGGGACGTGAACGCCGAGCCGAGACTCTTGGCTGCTGTAGCCTGTGC	434
OY	361	GGGGGGGCTCCGCGAGGCTTTTGGCTACCAAGGCGCGAATATGTATTAAGCCGCTGCC	420
Db	435	GGGGGGGCTCCGCGAGGCTTTTGGCTACCAAGGCGCGAATATGTATTAAGCCGCTGCC	494
OY	421	AATCTTAGCGCGCCGCGCGCGACAGCCAGACGCGAGCGACACCTTCTCAGCGCCGG	480
Db	495	AATCTTAGCGCGCGCGCGCGACAGCCAGACGCGAGCGACACCTTCTCAGCGCGGG	554
OY	481	GGTGTCCGGGGCGGCGCTTCCGAGAGACCCACCTCTCGCTGGGGGGTGGCTCGGGCTGG	540
Db	555	GGTGTCCGGGGCGGCGCTTCCGAGAGACCCACCTCTCGCTGGGGGGTGGCTCGGGCTGG	614
OY	541	AAACCCGCGCATCTTAAGGGGCGCTTGAACCTTACAGACCGCGGGCGGGGCTTGGGGAG	600
Db	615	AAACCCGCGCATCTTAAGGGGCGCTTGAACCTTACAGACCGCGGGCGGGGCTTGGGGAG	674
OY	601	AGTGTAGCCGCGCGAGCTGTGGGGCGCGCAAGCGTTGTGCTATACCGGCGAGTGTG	660
Db	675	AGTGTAGCCGCGCGAGCTGTGGGGCGCGCAAGCGTTGTGCTATACCGGCGAGTGTG	734
OY	661	GAGAGCGTGTGTGTGCGCGGAGAGTAAATGTCTACAGGCGCGGACCTGTGAACAT	720
Db	735	GAGAGCGTGTGTGTGCGCGGAGAGTAAATGTCTACAGGCGCGGACCTGTGAACAT	794
OY	721	TATCTGCTAGACGTGTGTGCGCAACGGCGGGCGGCACTTACCGCCATCCAGCATCTCTAAC	780
Db	795	TATCTGCTAGACGTGTGTGCGCAACGGCGGGCGGCACTTACCGCCATCCAGCATCTCTAAC	854
OY	781	CCCAATCAACATCGTTTGATGACAGAGGTGCTCTTAAAGATGAGACCTCGGGGCCAAG	840
Db	855	CCCAATCAACATCGTTTGATGACAGAGGTGCTCTTAAAGATGAGACCTCGGGGCCAAG	914
OY	841	GTCAACCGGCAATGCGGCCCTGAGCGCTGTCCGCAACTTGTGCTCGGCGAGAGAACTGAAC	900
Db	915	GTCAACCGGCAATGCGGCCCTGAGCGCTGTCCGCAACTTGTGCTCGGCGAGAGAACTGAAC	974
OY	901	AAATGAGTGAACAGACACCCCGAGTACTGGGAACTGTCCATCTCTTACCCAGCGAGGAC	960
Db	975	AAATGAGTGAACAGACACCCCGAGTACTGGGAACTGTCCATCTCTTACCCAGCGAGGAC	1034
OY	961	CTGTGTGGAGCACACCACTGTGAACACCGTGGGATGGCTGATGTGGGTACCTGTGTATAC	1020
Db	1035	CTGTGTGGAGCACACCACTGTGAACACCGTGGGATGGCTGATGTGGGTACCTGTGTATAC	1094
OY	1021	CCCAAGAGAACTCTGTGTATTGAGAGCATGTGGCTTCATCAGCTTCAACCACTGCC	1080
Db	1095	CCCAAGAGAACTCTGTGTATTGAGAGCATGTGGCTTCATCAGCTTCAACCACTGCC	1154
OY	1081	CACAGAGCGGGCCACGTTCAACATGCCCCCATGACAAATGTAAAGTCTGTAGAGAGTG	1140
Db	1155	CACAGAGCGGGCCACGTTCAACATGCCCCCATGACAAATGTAAAGTCTGTAGAGAGTG	1214
OY	1141	TTTGGGAAAGCTCCAGGCAACACCATGTATGCCCAACCCCTATCCAGATGACGCGTGC	1200
Db	1215	TTTGGGAAAGCTCCAGGCAACACCATGTATGCCCAACCCCTATCCAGATGACGCGTGC	1274
OY	1201	AAACCCGTGTAGCTCTGAGTGTGCTGCATATACCGACTTCTGTGACAGCGGGTACGGT	1260
Db	1275	AAACCCGTGTAGCTCTGAGTGTGCTGCATATACCGACTTCTGTGACAGCGGGTACGGT	1334

QY	1261	GAGGCTCTCGTGGACCAACCCAGACAGCCATCTCCGTCGCCGAGATCTCCGGGGCC	1320
Db	1335	GACTGCTCTCGGACCAACCCAGACAGCCATCTCCGTCGCCGAGATCTCCGGGGCC	1394
QY	1321	AGCTACACCCCTGAGCCAGAGAGTGCAGAGCGGGCTTTTGGCGTGGGCTCCAGCCCTGTCT	1380
Db	1395	AGCTACACCCCTGAGCCAGAGAGTGCAGAGCGGGCTTTTGGCGTGGGCTCCAGCCCTGTCT	1454
QY	1381	TACATGCATGACTGCAACCAAGCTGTGTGCACCCGGGAAAGGCCAAGGACAGATGTGTGC	1440
Db	1455	TACATGCATGACTGCAACCAAGCTGTGTGCACCCGGGAAAGGCCAAGGACAGATGTGTGC	1514
QY	1441	CAGACCCCGCACTTCCCTCGGGCCGATGGGACACACTGTGTGGAGAGGCAACTCTGCTTC	1500
Db	1515	CAGACCCCGCACTTCCCTCGGGCCGATGGGACACACTGTGTGGAGAGGCAACTCTGCTTC	1574
QY	1501	AAAGGGGCGCTCGCTGGAGAGACACAACCTCAACAAGCAGAGGATGATGATCTCTGGGCGC	1560
Db	1575	AAAGGGGCGCTCGCTGGAGAGACACAACCTCAACAAGCAGAGGATGATGATCTCTGGGCGC	1634
QY	1561	AAATGGATCCCATATGAGCCCTGTGCGCAGCATGTGTGGGGGGCTGACACTGGCAGG	1620
Db	1635	AAATGGATCCCATATGAGCCCTGTGCGCAGCATGTGTGGGGGGCTGACACTGGCAGG	1694
QY	1621	AGGCAATGCAACCAACCCCACTGGCAACGGGGGCAAGTACTGTGAGAGAGTGAAGGATG	1680
Db	1695	AGGCAATGCAACCAACCCCACTGGCAACGGGGGCAAGTACTGTGAGAGAGTGAAGGATG	1754
QY	1681	AAATACCGATCTCGCATCTGAGAGCCCTGCCAGTCAGCTCCGGAAAGAGGCTCCGG	1740
Db	1755	AAATACCGATCTCGCATCTGAGAGCCCTGCCAGTCAGCTCCGGAAAGAGGCTCCGG	1814
QY	1741	GAGGAGCAGTGTGAGGCTTTCAACGGCTCAACCAACAGACCAACACGGCTCACTCTGCC	1800
Db	1815	GAGGAGCAGTGTGAGGCTTTCAACGGCTCAACCAACAGACCAACACGGCTCACTCTGCC	1874
QY	1801	GTGGCATAGGGGTGCCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTCAACTATCTGC	1860
Db	1875	GTGGCATAGGGGTGCCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTCAACTATCTGC	1934
QY	1861	CGAGCCAAATGGCACTGGGCTCTCTATGTGTGCTGGGCAACCA---GGTGGTGGAGGGCAG	1917
Db	1935	CGAGCCAAATGGCACTGGGCTCTCTATGTGTGCTGGGCAACCAAGTGGATCAAGGCTGGCT	1994
QY	1918	CTGTGCTCTCTGACTCCACCTCCGCTGTGTGCCAAGGCAAGTGCATCAAGGCTGGCTGT	1977
Db	1995	CTGTGCTCTCTGACTCCACCTCCGCTGTGTGCCAAGGCAAGTGCATCAAGGCTGGCTGT	2054
QY	1978	GATGGGAACTGGGGCTCTCAAGAAAGATTCGACAGTGTGGGGTGTGTGGGGAGACAT	2037
Db	2055	GATGGGAACTGGGGCTCTCAAGAAAGATTCGACAGTGTGGGGTGTGTGGGGAGACAT	2114
QY	2038	AAGAGCTGCAAGAAAGGAGTGGACGCTTCACCAAGGCCATGATGGCTACAAATTTGSG	2097
Db	2115	AAGAGCTGCAAGAAAGGAGTGGACGCTTCACCAAGGCCATGATGGCTACAAATTTGSG	2174
QY	2098	GTGGCACTCCCGCAGCGGCTCAACCATTCGACATCCGCAAGCGGATTAACAAAGGCGTG	2157
Db	2175	GTGGCACTCCCGCAGCGGCTCAACCATTCGACATCCGCAAGCGGATTAACAAAGGCGTG	2254
QY	2158	ATCGGGGATGACACTACTCTGGCTGTGAAGAAACGCCAAGGCAAGTACTCTCTCAACGGG	2217
Db	2235	ATCGGGGATGACACTACTCTGGCTGTGAAGAAACGCCAAGGCAAGTACTCTCTCAACGGG	2294
QY	2218	CATTTCGTGGTGTGGGGGGGAGGCGGGACGCTGGTGTGAAGGGAGTCTCTCTGGGTAC	2277
Db	2295	CATTTCGTGGTGTGGGGGGGAGGCGGGACGCTGGTGTGAAGGGAGTCTCTCTGGGTAC	2354
QY	2278	AGCGGCAAGGGCAAGGGGTGAGAGACCTCGACAGGCTCCCGGCCATCTCTGAGAGCGGTG	2337
Db	2355	AGCGGCAAGGGCAAGGGGTGAGAGACCTCGACAGGCTCCCGGCCATCTCTGAGAGCGGTG	2414
QY	2338	ACCGTGAAGTCTCTCCGTGGGGGAAGATGACACCGCCCGGGGTCCGCTACTCTCTTAT	2397

```

Db      2415  ACCGTTGAGGCTCTCTCCGTGGGGAAGATGACACGCCCCGGGCTCCGCTACTCTTCTAT 2474
QY      2398  CTGCCCAAGAGCCTCGGAGAGACAAGTCCCTCATCCC 2436
          |||||||||||||||||||
Db      2475  CTGCCCAAGAGCCTCGGAGAGACAAGTCCCTCATCCC 2513

RESULT 6
US-10-163-316-3
; Sequence 3, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65352, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-10-163-316-3

Query Match          52.3%; Score 1491; DB 41; Length 2465;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGCTTCTGCTGGGCACTCTAACCTTGCTTTCGCGCGGCAACCGCTGGAGCTCTAG 60
Db      1  ATGCTTCTGCTGGGCACTCTAACCTTGCTTTCGCGCGGCAACCGCTGGAGCTCTAG 60
QY      61  CCAGAGCGGAGAGTAGTCGTTCCATCCGACTGGACCCGGACATTAAACGGCGCCGTAC 120
Db      61  CCAGAGCGGAGAGTAGTCGTTCCATCCGACTGGACCCGGACATTAAACGGCGCCGTAC 120
QY      121  TACTGGCGGGGTCCCGAGGACTCCGGGGATCAGGAGCTATTTTCAGATCAGACATT 180
Db      121  TACTGGCGGGGTCCCGAGGACTCCGGGGATCAGGAGCTATTTTCAGATCAGACATT 180
QY      181  CAGAGGACTTTTACCTACACCTGACGCGCGAGATGCTCATTTCTTGCTCCGCCCTTCC 240
Db      181  CAGAGGACTTTTACCTACACCTGACGCGCGAGATGCTCATTTCTTGCTCCGCCCTTCC 240
QY      241  ACTGAGCATCTGGGCGTCCGCCCTTCAGGGGCTCACCGGGGGCTTTCAGACTGCGACG 300
Db      241  ACTGAGCATCTGGGCGTCCGCCCTTCAGGGGCTCACCGGGGGCTTTCAGACTGCGACG 300
QY      301  TGCCTTATTCTTGGGGAGCTGAACGCCGAGACCGGACTCGTCTGCTGTGTGACCTTGG 360
Db      301  TGCCTTATTCTTGGGGAGCTGAACGCCGAGACCGGACTCGTCTGCTGTGTGAGCTTGG 360
QY      361  GGGGGGCTCCGCGAGGCTTTGGCTACGAGCGCGGAGTATGTCAATTAAGCCGCTGCC 420
Db      361  GGGGGGCTCCGCGAGGCTTTGGCTACGAGCGCGGAGTATGTCAATTAAGCCGCTGCC 420
QY      421  AATGCTAGCGGCGCGGGGCGAGCGAAGCAGCGAGGGGCGACACCTTCCAGCGCGG 480
Db      421  AATGCTAGCGGCGCGGGGCGAGCGAAGCAGCGAGGGGCGACACCTTTCAGAGCGCGG 480
QY      481  GGTGTTCCGGGCGGCTTCGCGAGACCCACCTCTCGCTGGGGGGTGCGCTCGGGCTGG 540
Db      481  GGTGTTCCGGGCGGCTTCGCGAGACCCACCTCTCGCTGGGGGGTGCGCTCGGGCTGG 540
QY      541  AAGCCGCGATCTTACGGGCGCTTGAGCCTTACAGCGCGGCGGGGCTTGGGGAG 600
          |||||||||||||||||||

```



```

Db 541 AACCCGGCATTCTACGGGCCCCCTGACCCCTTACAAGCCGCGGCGGGCTTGGGGAG 600
QY 601 AGTCGTACGCGGCGAGGCTCTGGGGCGCCCAAGCGTTTCGTCTATCCCGGCTACCTG 660
Db 601 AGTCGTACGCGGCGAGGCTCTGGGGCGCCCAAGCGTTTCGTCTATCCCGGCTACCTG 660
QY 661 GAGAGCGTGGTGGCGGAGCATCATGTGCAAGTTCCACGGCGCGGAGCTGGAACAT 720
Db 661 GAGAGCGTGGTGGCGGAGCATCATGTGCAAGTTCCACGGCGCGGAGCTGGAACAT 720
QY 721 TATCTGTAGAGCTGTGGCAAGCGGCGGAGCTCTACCGCCAMCCAGCATCTCAAC 780
Db 721 TATCTGTAGAGCTGTGGCAAGCGGCGGAGCTCTACCGCCAMCCAGCATCTCAAC 780
QY 781 CCCATCAACATCGTTGTGTGTCAGAGTGTCTTCTTAGAGATCGTGAATCCGGGCCAAG 840
Db 781 CCCATCAACATCGTTGTGTGTCAGAGTGTCTTCTTAGAGATCGTGAATCCGGGCCAAG 840
QY 841 GTCACCGGCAATGCGGCCCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAGACTGAAC 900
Db 841 GTCACCGGCAATGCGGCCCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAGACTGAAC 900
QY 901 AAAGTAGTGAACAAGCACCAGAGTACTGGGACATGCGCATCTTCCACAGGCGAGAC 960
Db 901 AAAGTAGTGAACAAGCACCAGAGTACTGGGACATGCGCATCTTCCACAGGCGAGAC 960
QY 961 CTGTGTGAGGCAACCACTGTGACACCCCTGGGCAATGGCTGATGTGGGTACCATGTGTAC 1020
Db 961 CTGTGTGAGGCAACCACTGTGACACCCCTGGGCAATGGCTGATGTGGGTACCATGTGTAC 1020
QY 1021 CCCAAGAAACCTGCTCTGTCTATGTAGAGAGATGGGCTTCCATGACCTTCACCACTGCC 1080
Db 1021 CCCAAGAAACCTGCTCTGTCTATGTAGAGAGATGGGCTTCCATGACCTTCACCACTGCC 1080
QY 1081 CACGAGCTGGGCGCAAGTTCATCATGATGCCCCATGATGAAAGTCTGTGTGAGAGAGTG 1140
Db 1081 CACGAGCTGGGCGCAAGTTCATCATGATGCCCCATGATGAAAGTCTGTGTGAGAGAGTG 1140
QY 1141 TTTGGGAAGCTCGAGCCCAACCATGATGTCCCGAACCCTCATCAGATCGAACCTGTGC 1200
Db 1141 TTTGGGAAGCTCGAGCCCAACCATGATGTCCCGAACCCTCATCAGATCGAACCTGTGC 1200
QY 1201 AACCCGTGACGCTGTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 AACCCGTGACGCTGTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 GACTGCTCTGTGACCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTGTGACCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCTGTGACCAAGCAAGTGTGGCTTTTGGCGTGGGCTTCCAAAGCCCTGTCT 1380
Db 1321 AGCTACACCTGTGACCAAGCAAGTGTGGCTTTTGGCGTGGGCTTCCAAAGCCCTGTCT 1380
QY 1381 TACATGAGTGTGCTGACCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1381 TACATGAGTGTGCTGACCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 1440
QY 1441 CAGACCGGCAATTCCTGCGGCGATGATGACACAGCTGTGGGAGGCAAGCTGTGCTG 1500
Db 1441 CAGACCGGCAATTCCTGCGGCGATGATGACACAGCTGTGGGAGGCAAGCTGTGCTG 1500
QY 1501 AAAGGGGCTGTGCTGAGAGACACACCTCAACAGCAGAG 1542
Db 1501 AAAGGGGCTGTGCTGAGAGACACACCTCAACAGCAGAG 1542

```

RESULT 7

```

US-60-297-863-3
; Sequence 3, Application US/60297863
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses

```

```

; TITLE OF INVENTION: therefor
; FILE REFERENCE: MPI2001-025PI(M)
; CURRENT APPLICATION NUMBER: US/60/297,863
; CURRENT FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-60-297-863-3

Query Match          52.3%; Score 1491; DB 73; Length 2469;
Best local similarity 99.9%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCAATCCCTTAACCCCTTTCGCGCGGCGAAGCGCTGAGAGCTGTAG 60
Db 1 ATGCTTCTGCTGGGCAATCCCTTAACCCCTTTCGCGCGGCGAAGCGCTGAGAGCTGTAG 60
QY 61 CCAGAGCGGAGAGTACTGCTTCCATCCGACTGAGACCGGACATTACGCGCGCTAC 120
Db 61 CCAGAGCGGAGAGTACTGCTTCCATCCGACTGAGACCGGACATTACGCGCGCTAC 120
QY 121 TACTGCGGGGTCGCGAGGAGTCCGCGGATGTCAGGAGTCAATTTTACATACAGCATTT 180
Db 121 TACTGCGGGGTCGCGAGGAGTCCGCGGATGTCAGGAGTCAATTTTACATACAGCATTT 180
QY 181 CAGGAGACTTTTACTACCTGACGCGCGAGTGTCTGATTTTGGCTCCGCTCTCC 240
Db 181 CAGGAGACTTTTACTACCTGACGCGCGAGTGTCTGATTTTGGCTCCGCTCTCC 240
QY 241 ACTGAGACTGTGGGCTCCCCCTCAAGGGGCTACCGGGGCTCTTCAGACCTGACGC 300
Db 241 ACTGAGACTGTGGGCTCCCCCTCAAGGGGCTACCGGGGCTCTTCAGACCTGACGC 300
QY 301 TGTCTTATTTTGGGAGCTGAAAGCGCGAGCGGAGCTGTTGGCTGTGAGCTGTGC 360
Db 301 TGTCTTATTTTGGGAGCTGAAAGCGCGAGCGGAGCTGTTGGCTGTGAGCTGTGC 360
QY 361 GGGGGGCTCCGCGAGCCCTTGGCTACCGAGGCGCGAGTATGATTAAGCCCGTCC 420
Db 361 GGGGGGCTCCGCGAGCCCTTGGCTACCGAGGCGCGAGTATGATTAAGCCCGTCC 420
QY 421 AATGCTAGGCGCGCGGCGGCGAGCGCAACAGCCAGGGCGCACACCTTCTCCAGCGCG 480
Db 421 AATGCTAGGCGCGCGGCGGCGAGCGCAACAGCCAGGGCGCACACCTTCTCCAGCGCG 480
QY 481 GGTGTTCCGGGGGGGCTTCCGAGAGCCCACTCTGCGTGGGGGGTGGGCTGG 540
Db 481 GGTGTTCCGGGGGGGCTTCCGAGAGCCCACTCTGCGTGGGGGGTGGGCTGG 540
QY 541 AACCCCGCATCTCAAGGCGCTTGAACCCCTTCAAGCCCGGCGGCGGCTTGGGAG 600
Db 541 AACCCCGCATCTCAAGGCGCTTGAACCCCTTCAAGCCCGGCGGCGGCTTGGGAG 600
QY 601 AGTCGTAGCGGCGAGGCTGTGGGCGCGCAAGCGTTTCTGTATCTCCCGGTAGCTG 660
Db 601 AGTCGTAGCGGCGAGGCTGTGGGCGCGCAAGCGTTTCTGTATCTCCCGGTAGCTG 660
QY 661 GAGAGCTGTGTTGGGAGAGTCAATGTCAGATTCCAGGGGCGGAGCTGGAACAT 720
Db 661 GAGAGCTGTGTTGGGAGAGTCAATGTCAGATTCCAGGGGCGGAGCTGGAACAT 720
QY 721 TATCTGTAGAGCTGTGGAACCGGCGGAGCTTCAACCGCATATCCAGCATCTCAAC 780
Db 721 TATCTGTAGAGCTGTGGAACCGGCGGAGCTTCAACCGCATATCCAGCATCTCAAC 780
QY 781 CCCATCAACATCGTTGTGTGTCAGAGTGTCTTCTTAGAGATCGTGAATCCGGGCCAAG 840
Db 781 CCCATCAACATCGTTGTGTGTCAGAGTGTCTTCTTAGAGATCGTGAATCCGGGCCAAG 840

```

```

Db 781 CCATCAACATCGTGTGTGCAAGTGCTGCTTCTTAGA..ATGTGATCCGCGCCCAAG 840
QY 841 GTCAACGGCAATGGGGCCCTGAGCTGCGGCACTTCTGTCCTCGGAGAAAGACGTCGAC 900
Db 841 GTCAACGGCAATGGGGCCCTGAGCTGCGGCACTTCTGTCCTCGGAGAAAGACGTCGAC 900
QY 901 AAAGTAGTGACAAAGCAACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 960
Db 901 AAAGTAGTGACAAAGCAACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 960
QY 961 CTGTGTGAGACCAACCACTGTGACACCCCTGGCATGGCT..ATGTGGGTACATGTGTGAC 1020
Db 961 CTGTGTGAGACCAACCACTGTGACACCCCTGGCATGGCT..ATGTGGGTACATGTGTGAC 1020
QY 1021 CCCAAGAGAAGTGTCTGTCTGTCATTGAGGAGATGGGCTTCATCAGCCTTTCACCATGTC 1080
Db 1021 CCCAAGAGAAGTGTCTGTCTGTCATTGAGGAGATGGGCTTCATCAGCCTTTCACCATGTC 1080
QY 1081 CACGAGCTGGGCGACGTGTTCACATGCCCCATGACAAATGTGAAAGTCTGTGAGAGGTG 1140
Db 1081 CACGAGCTGGGCGACGTGTTCACATGCCCCATGACAAATGTGAAAGTCTGTGAGAGGTG 1140
QY 1141 TTGGGAGAGTCCGAGCAACCAACATGATGTGCCCCGACCTCTATCCAGATCGACCTGCC 1200
Db 1141 TTGGGAGAGTCCGAGCAACCAACATGATGTGCCCCGACCTCTATCCAGATCGACCTGCC 1200
QY 1201 AACCCCTGTGACGCTGTGACATGCTGCTCATATCAGCAGCTTCTCTGAGACAGCGGACG 1260
Db 1201 AACCCCTGTGACGCTGTGACATGCTGCTCATATCAGCAGCTTCTCTGAGACAGCGGACG 1260
QY 1261 GATGCTCTCTGTGACCAACCCAGCAAGCCATCTCCCTGCCGAGATCTGCGGGCGCC 1320
Db 1261 GATGCTCTCTGTGACCAACCCAGCAAGCCATCTCCCTGCCGAGATCTGCGGGCGCC 1320
QY 1321 AGCTACACCTGTGAGCAGAGTGAGTGAGGCTTTGGGGTGGGCTCCAGAGCCCTGTCT 1380
Db 1321 AGCTACACCTGTGAGCAGAGTGAGTGAGGCTTTGGGGTGGGCTCCAGAGCCCTGTCT 1380
QY 1381 TACATCAGTACTGTGACCAAGCTGTGTGTGACCGGAGAGGCCAAGGACATGTGTGTC 1440
Db 1381 TACATCAGTACTGTGACCAAGCTGTGTGTGACCGGAGAGGCCAAGGACATGTGTGTC 1440
QY 1441 CACACCCGCGCACTTCCCTGGGCGGATGGGACGACTGTGTGGAGGCAAGCTGTGCTC 1500
Db 1441 CACACCCGCGCACTTCCCTGGGCGGATGGGACGACTGTGTGGAGGCAAGCTGTGCTC 1500
QY 1501 AAAGGGGCTGTGAGAGACACAACTCAACAAAGCAGG 1542
Db 1501 AAAGGGGCTGTGAGAGACACAACTCAACAAAGCAGG 1542

RESULT 8
US-10-163-316-1
; Sequence 1, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:

```

```

; NAME/KEY: misc.feature
; LOCATION: (1)-(2940)
; OTHER INFORMATION: n = A,T,C or G
US-10-163-316-1

Query Match
Best Local Similarity 99.9%, Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGTGGGCACTTAACCTGCTTTCCGCGGCAACCGCTGAGGCTGTGAG 60
Db 472 ATGCTTCTGTGGGCACTTAACCTGCTTTCCGCGGCAACCGCTGAGGCTGTGAG 531
QY 61 CCAGAGCGGAGGTGCTGCCATCCGATGAGACCCGGACATTAAAGCGCCGCTAC 120
Db 532 CCAGAGCGGAGGTGCTGCCATCCGATGAGACCCGGACATTAAAGCGCCGCTAC 591
QY 121 TACTGGCGGGGTCCGAGAGACTCCGGGATCAGGACATTTTTCAGATCAGACATTT 180
Db 592 TACTGGCGGGGTCCGAGAGACTCCGGGATCAGGACATTTTTCAGATCAGACATTT 651
QY 181 CAGAGGACTTTTACCTACACCTGACGCGGAGATGCTAGTTCTTGCTCCGCTTCC 240
Db 652 CAGAGGACTTTTACCTACACCTGACGCGGAGATGCTAGTTCTTGCTCCGCTTCC 711
QY 241 ACTGAGCATTTGGGGCTCCCTCCAGGGGCTCACCGGGGGCTCTTGACAGCTGACGC 300
Db 712 ACTGAGCATTTGGGGCTCCCTCCAGGGGCTCACCGGGGGCTCTTGACAGCTGACGC 771
QY 301 TGCCTTCTATTTGGGAGGTAGAGCCGAGCGGACTGTTGCTGCTGTGAGCTGTGC 360
Db 772 TGCCTTCTATTTGGGAGGTAGAGCCGAGCGGACTGTTGCTGCTGTGAGCTGTGC 831
QY 361 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCGCTGCC 420
Db 832 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCGCTGCC 891
QY 421 AATGTACGCGCGCGGCGGCGCAGCGCAACAGCGAGGCGGACACTTCTACAGCGCGG 480
Db 892 AATGTACGCGCGCGGCGGCGCAGCGCAACAGCGAGGCGGACACTTCTACAGCGCGG 951
QY 481 GGTGTTCCGGGCGGGCTTTCGGAAGCCACCTCTGCTGCGGGGTGGCTCGGGCTGG 540
Db 952 GGTGTTCCGGGCGGGCTTTCGGAAGCCACCTCTGCTGCGGGGTGGCTCGGGCTGG 1011
QY 541 AACCCCGCATCTACAGGCGCCCTTGACCTTACAAAGCGCGGGCGGCTTGGGGAG 600
Db 1012 AACCCCGCATCTACAGGCGCCCTTGACCTTACAAAGCGCGGGCGGCTTGGGGAG 1071
QY 601 AGCTGTAGCCGGCGAGGCTGTGGGCGGCGCAAGGCTTTCGTTATCCCGGCTACGTG 660
Db 1072 AGCTGTAGCCGGCGAGGCTGTGGGCGGCGCAAGGCTTTCGTTATCCCGGCTACGTG 1131
QY 721 TATCTGTGAGGCTGTGGGCAACGGCGGGGAGATCTACAGGCACTCCAGATCTTAA 780
Db 1192 TATCTGTGAGGCTGTGGGCAACGGCGGGGAGATCTACAGGCACTCCAGATCTTAA 1251
QY 781 CCCATCAACATCGTTGTGTCAAGTCTGCTTTCAGAGATGCTGACTCCGGGCGCAAG 840
Db 1252 CCCATCAACATCGTTGTGTCAAGTCTGCTTTCAGAGATGCTGACTCCGGGCGCAAG 1311
QY 841 GTCAACGGCAATGGGCGCTGTGACGCTGCGCAACTTGTGTCTGTGCGAGAAAGCTGAC 900
Db 1312 GTCAACGGCAATGGGCGCTGTGACGCTGCGCAACTTGTGTCTGTGCGAGAAAGCTGAC 1371
QY 901 AAAGTAGTGACAAAGCAACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 960
Db 1372 AAAGTAGTGACAAAGCAACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 1431

```



```
Db 1612 TTTGGGAGGCTCCGAGGCAACCAATGATGTCCCGACCCCTCATCCAGATCGACCTGGCC 1671
QY 1201 AACCCCTGTCAGCTGTCAGTGTGTCATCATCAGCAGCTCTCTGACAGCGGGGCGCT 1260
Db 1672 AACCCCTGGTCAGCTGTCAGTGTGTCATCATCAGCAGCTCTCTGACAGCGGGGCGCT 1731
QY 1261 GACTGCTCTGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1320
Db 1732 GACTGCTCTGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1791
QY 1321 AGCTACACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1380
Db 1792 AGCTACACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1851
QY 1381 TACATGAGTACTGACACCAAGCTGTGTGTCACCGGAGGAGGAGCAGATGGTGTGC 1440
Db 1852 TACATGAGTACTGACACCAAGCTGTGTGTCACCGGAGGAGGAGCAGATGGTGTGC 1911
QY 1441 CAGACCCGCGCACTTCCCTGGGCGCATGGACACGCTGTGGGAGGAGGAGCAGTGTGC 1500
Db 1912 CAGACCCGCGCACTTCCCTGGGCGCATGGACACGCTGTGGGAGGAGGAGCAGTGTGC 1971
QY 1501 AAGGGGCGCTGCTGGAGAGACCAACCTCACAACAGCAGG 1542
Db 1972 AAGGGGCGCTGCTGGAGAGACCAACCTCACAACAGCAGG 2013

RESULT 10
US-10-093-463-27
Sequence 27, Application US/10093463
GENERAL INFORMATION:
APPLICANT: Padigar, Muralidhara
APPLICANT: Shenoy, Suresh
APPLICANT: Kekuda, Ramesh
APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal
APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glenda
APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Zernusen, Bryan
APPLICANT: Tcherny, Velizar
APPLICANT: Gangoli, Esha
APPLICANT: Vernet, Corine
APPLICANT: Pena, Carol
APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patturajan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
FILE REFERENCE: 21402-290A (Cura 590A)
CURRENT FILING DATE: 2002-06-24
PRIOR FILING DATE: 2002-06-24
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
```

```

PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 2804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (25)..(2797)
US-10-093-463-27

Query Match 52.2% Score 1488; DB 39; Length 2804;
Best Local Similarly 99.9%; Pred. No. 0;
Matches 1538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCGCTGGGATCTTAACCTCTGCTTCCCGGGGAGACCGCTGAGCTCTGAG 60
Db 25 ATGCTTCGCTGGGATCTTAACCTCTGCTTCCCGGGGAGACCGCTGAGCTCTGAG 84
QY 61 CCAGAGCGGAGTAGTGTGTTCCCATCGACTGAGCCCGGACATTACGGCGCGGTAC 120
Db 85 CCAGAGCGGAGTAGTGTGTTCCCATCGACTGAGCCCGGACATTACGGCGCGGTAC 144
QY 121 TACTGGCGGGGTCGCGAGGACTCGGGGATCAGGACTATTTTTCAGATCAGCATTT 180
Db 145 TACTGGCGGGGTCGCGAGGACTCGGGGATCAGGACTATTTTTCAGATCAGCATTT 204
QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTGCTGCTCCCGCTTCTCC 240
Db 205 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTGCTGCTCCCGCTTCTCC 264
QY 241 ACTGAGCATCTGGGCTGCCCTTCACAGGGGCTCACCGGGGCTCTTCAGACTGCGAGC 300
Db 265 ACTGAGCATCTGGGCTGCCCTTCACAGGGGCTCACCGGGGCTCTTCAGACTGCGAGC 324
QY 301 TGTCTTATCTTGGGAGCGTGAACCGGAGCGGACTCTTGGCTGTGAGCTGTGC 360
Db 325 TGTCTTATCTTGGGAGCGTGAACCGGAGCGGACTCTTGGCTGTGAGCTGTGC 384
QY 361 GGGGGCTCCGCGGAGCTTGGCTACCGAGGCGCGAGATGCTTACCGCGGTGCGCC 420
Db 385 GGGGGCTCCGCGGAGCTTGGCTACCGAGGCGCGAGATGCTTACCGCGGTGCGCC 444
```

QY 421 AATGCTAGCGGCGCGGCGGCGAGCGCAAGCCAGCGGCGCACACCTTCUCCAGCGCGG 480
DB 445 AATGCTAGCGGCGCGGCGGCGGCGAGCGCAAGCCAGCGGCGCACACCTTCUCCAGCGCGG 504
QY 481 GGTGTTCCGGGCGGCGCTTCGCGAGAGCCCACTCTGCTGCGCGGCGGCTGCGGCGTGG 540
DB 505 GGTGTTCCGGGCGGCGCTTCGCGAGAGCCCACTCTGCTGCGGCGGCGGCTGCGGCGTGG 564
QY 541 AATGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 565 AATGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 624
QY 601 AGTGTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 625 AGTGTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 684
QY 661 GAGACGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 685 GAGACGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 744
QY 721 TATGCTGAGCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 745 TATGCTGAGCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 804
QY 781 CCCATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 805 CCCATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864
QY 841 GTCACCGGCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 865 GTCACCGGCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 924
QY 901 AAATGAGTGCAAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 925 AAATGAGTGCAAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 984
QY 961 CTGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 985 CTGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1044
QY 1021 CCCAAGAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1045 CCCAAGAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104
QY 1081 CACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1105 CACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1164
QY 1141 TTTGGGAAGCTCCGAGCGCAACACATGATGCTCCCGACCTTCATCCAGATCGACCGTGGC 1200
DB 1165 TTTGGGAAGCTCCGAGCGCAACACATGATGCTCCCGACCTTCATCCAGATCGACCGTGGC 1224
QY 1201 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB 1225 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1284
QY 1261 GACTGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
DB 1285 GACTGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1344
QY 1321 AGCTACACCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1345 AGCTACACCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1404
QY 1381 TACATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1405 TACATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1464
QY 1441 CAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
DB 1465 CAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1524
QY 1501 AAAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1539

DB 1525 AAAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1563

|||||

RESULT 11
US-10-093-463-29
Sequence 29, Application US/10093463
GENERAL INFORMATION:
APPLICANT: Padigar, Muralidhara
APPLICANT: Shenoy, Suresh
APPLICANT: Kerkuda, Ramesh
APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal
APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glenda
APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Berenc
APPLICANT: Li, Li
APPLICANT: Zernusen, Bryan
APPLICANT: Tcherny, Velizar
APPLICANT: Gangolli, Esha
APPLICANT: Vernet, Corine
APPLICANT: Pena, Carol
APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patlura, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic
FILE REFERENCE: Encoding the Antigens, and Methods of Use.
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: US/10/093,463
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18

|||||
Db 2269 CTCGGGAAGAGCTTCGGGAGAGACGTGAGGCTTCAACGGGTACACACAGCAC 2328
1782 CAACGGGCTACTCTGCGGTGAGATGGTGCCCAAGTACTCGGCGTGTCTCCCGGGA 1841
2329 CAACCGGCTACTCTGCGGTGAGATGGTGCCCAAGTACTCGGCGTGTCTCCCGGGA 2388
1842 CAAGTCAAGCTATCTGCGGAGACCAATGAGCACTGCTCTTCTATGTCGTGACCCCA 1901
2389 CAAGTCAAGCTATCTGCGGAGACCAATGAGCACTGCTCTTCTATGTCGTGACCCCA 2448
1902 GGTGTGAGACGACGCTGTCTCTCTGACTCCACTCCGTCTGTGTCTCAAGGCAATG 1961
2449 GGTGTGAGACGACGCTGTCTCTCTGACTCCACTCCGTCTGTGTCTCAAGGCAATG 2508
1962 CATCAAGGCTGCTGTGATGAGGAACTGGGCTCCAGAGAGATTCGACAACTGTGGGT 2021
2509 CATCAAGGCTGCTGTGATGAGGAACTGGGCTCCAGAGAGATTCGACAACTGTGGGT 2568
2022 GTGTGGGGAGACAAATAGAGCTGCAAGAGAGTGAATGACTCTTCCACCAAGCCATGCA 2081
2569 GTGTGGGGAGACAAATAGAGCTGCAAGAGAGTGAATGACTCTTCCACCAAGCCATGCA 2628
2082 TGGCTACAAATTTGTGTGGCCATCCCGAGGCGCTCAAGCATGACATCCGCGAGCG 2141
2629 TGGCTACAAATTTGTGTGGCCATCCCGAGGCGCTCAAGCATGACATCCGCGAGCG 2688
2142 CGGTTCAAAAGGGCTGATCGGGGATGACAACTAGGCTGTCAGAACACCAAGCAAGCAA 2201
2689 CGGTTCAAAAGGGCTGATCGGGGATGACAACTAGGCTGTCAGAACACCAAGCAAGCAA 2748
2202 GTACTGCTCAACGGGCAATTCGTGTGTGTGCGGGTGGAGACCGGACCTGTGTGAAGG 2261
2749 GTACTGCTCAACGGGCAATTCGTGTGTGTGCGGGTGGAGACCGGACCTGTGTGAAGG 2808
2262 CAGTCTGCTCGGTACAGCGGCAAGCGGCAAGCGGTGGAGAGCTTCAGGCTTCCGGGC 2321
2809 CAGTCTGCTCGGTACAGCGGCAAGCGGCAAGCGGTGGAGAGCTTCAGGCTTCCGGGC 2868
2322 CATCTTGGAGCGGCTGACGCTGGAGGTCCTCGTGGGGAGATGACACCGCGCGGT 2381
2869 CATCTTGGAGCGGCTGACGCTGGAGGTCCTCGTGGGGAGATGACACCGCGCGGT 2928
2382 CCGCTACTCTTATCTGTGCCAAAGAGCTGCGGAGAGACAACTCTCTATCCAGGA 2441
2929 CCGCTACTCTTATCTGTGCCAAAGAGCTGCGGAGAGACAACTCTCTATCCAGGA 2988
2442 CCGCGGGGAGCCCTCTGTCTTGGACAAACAGCTCTCAGCTCTCCAAACAGGTGAGGA 2501
2989 CCGCGGGGAGCCCTCTGTCTTGGACAAACAGCTCTCAGCTCTCCAAACAGGTGAGGA 3048
2502 GCGGAGAGAGCGCCCTCTGACGCTGGTGGTGGAGCTGGGGGCGGTCTCCCGGAG 2561
3049 GCGGAGAGAGCGCCCTCTGACGCTGGTGGTGGAGCTGGGGGCGGTCTCCCGGAG 3108
2562 CTGCGGAGTGGCTTGCAGAAAGCGGGGCTGTGACTGTCCGGGCTCCGCGGAGGCGAC 2621
3109 CTGCGGAGTGGCTTGCAGAAAGCGGGGCTGTGACTGTCCGGGCTCCGCGGAGGCGAC 3168
2622 GGTGCGGCTGCTGTGATGAGGCAATGAGGCTGCGGAGAGACCAAGCCGCGGGAGCCG 2681
3169 GGTGCGGCTGCTGTGATGAGGCAATGAGGCTGCGGAGAGACCAAGCCGCGGGAGCCG 3228
2682 CCGGAGTGGGAGCTGACGCTGTGACCCCTGCTCCAAAGAGTGGGCGCGGAGATTTCA 2741
3229 CCGGAGTGGGAGCTGACGCTGTGACCCCTGCTCCAAAGAGTGGGCGCGGAGATTTCA 3288
2742 GAGGCGCTCACTCAAGTGTGGGCGACGAGAGCGCGCTCTGCGCGGAGCAATGAGCA 2801
3289 GAGGCGCTCACTCAAGTGTGGGCGACGAGAGCGCGCTCTGCGCGGAGCAATGAGCA 3348
2802 CTTCACCGCAAGCCGAGAGCTGTGAGGCTTCTGTGAGGCGGTGCA 2853
|||||

Db 3349 CTTCACCGCAAGCCCGAGAGCTGACTTCTGCTGTGAGGCGCTGCTGA 3400
RESULT: 13
US-60-242-679-1814
; Sequence 1814, Application US/60242679
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steven Istvan
; APPLICANT: Spier, Eugene
; APPLICANT: Greenberg, Simon
; APPLICANT: Brandenberger, Ralph
; APPLICANT: Wang, Yu
; APPLICANT: Dudman, Alex
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000898-PROV
; CURRENT APPLICATION NUMBER: US/60/242,679
; NUMBER OF SEQ ID NOS: 2265
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1814
; LENGTH: 3759
; TYPE: DNA
; ORGANISM: HUMAN
US-60-242-679-1814
Query Match 44.2%; Score 1261; DB 68; Length 3759;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1542 GGTGATGATGCTCTGCGGCAATAGGATCCTTATGCGCCCTGCTGCGGCAATGTGTG 1601
2448 GGTGATGATGCTCTGCGGCAATAGGATCCTTATGCGCCCTGCTGCGGCAATGTGTG 2507
1602 GGGGCTGAGCTGGGCTGAGGAGGAGTGCACCAACCCCTGCTCAAGGAGGAGCAAGTA 1661
2508 GGGGCTGAGCTGGGCTGAGGAGGAGTGCACCAACCCCTGCTCAAGGAGGAGCAAGTA 2567
1662 CTGCGAGGAGTGAAGGTTGAATACCGATCTGCAATCTGAGAGCCCTGCGGCAAGTACG 1721
2568 CTGCGAGGAGTGAAGGTTGAATACCGATCTGCAATCTGAGAGCCCTGCGGCAAGTACG 2627
1722 CTCCGGAAGAGCTTCCGCGGAGAGAGTGTGAGGCTTTCAAGGCTTCAACACACACAC 1781
2628 CTCCGGAAGAGCTTCCGCGGAGAGAGTGTGAGGCTTTCAAGGCTTCAACACACACAC 2687
1782 CAACCGGCTCACTCTGCGCGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1841
2688 CAACCGGCTCACTCTGCGCGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2747
1842 CAAGTGAAGCTCATCTGCGGAGCCCAATGAGCACTGCTACTTCTATGTGTGTGCAACCA 1901
2748 CAAGTGAAGCTCATCTGCGGAGCCCAATGAGCACTGCTACTTCTATGTGTGTGCAACCA 2807
1902 GGTGTGAGAGCGGACGCTGTGCTCTGACCTCCAGCACTGCTGTGTGTGTGTGTGTGTGT 1961
2808 GGTGTGAGAGCGGACGCTGTGCTCTGACCTCCAGCACTGCTGTGTGTGTGTGTGTGTGT 2867
1962 CATCAAGGCTGCTGTGATGAGGCAATGAGGCTGCGGAGAGATTCGACAACTGTGGGT 2021
2868 CATCAAGGCTGCTGTGATGAGGCAATGAGGCTGCGGAGAGATTCGACAACTGTGGGT 2927
2022 GTGTGGGGAGACAAATAGAGCTGCAAGAGTGAATGACTCTTCCACCAAGCCATGCA 2081
2928 GTGTGGGGAGACAAATAGAGCTGCAAGAGTGAATGACTCTTCCACCAAGCCATGCA 2987
2082 TGGCTACAAATTTGTGTGGGCGACATCCCGAGGCGCTCAAGCATGACATCCGCGAGCG 2141
2988 TGGCTACAAATTTGTGTGGGCGACATCCCGAGGCGCTCAAGCATGACATCCGCGAGCG 3047
2142 CGTTTACAAAGGCTGATCGGGGATGACAACTACTGCTGTGAGAAACAGCAAGCA 2201
|||||

Db	3048	CGGTTACAAAGGCGTGMATCGGGGATGACAACTACCTGCGCTCTGTGAGAACACGCCAAAGGCA	3107
QY	2202	GTACCTGCTCAACGGGCAATTTCTGTGTGTGCGCGGTGGAGCGGGACCTGTGTGTGTAAGG	2261
Db	3108	GTACCTGCTCAACGGGCAATTTCTGTGTGTGCGCGGTGGAGCGGGACCTGTGTGTGTAAGG	3167
QY	2262	CAGTCTGCTGCGTACAGCGGCAACGGGGCAACAGGGGTGAGACCTGCAAGGCTTCCGGGC	2321
Db	3168	CAGTCTGCTGCGGTACAGCGGCAACGGGGCAACAGGGGTGAGACCTGCAAGGCTTCCGGGC	3227
QY	2322	CATCCTGGAGCGGCTGACCGGTGAGAGGTCTCTCCGTGGGGAGATGACACCGCCCGCGGT	2381
Db	3228	CATCCTGGAGCGGCTGACCGGTGAGAGGTCTCTCCGTGGGGAGATGACACCGCCCGCGGT	3287
QY	2382	CCGGTACCTCCTTCAATCTGTGCGCAAAAGCCCTGGGGAGGCAAAAGTCTCTATCCCAAG	2441
Db	3288	CCGGTACCTCCTTCAATCTGTGCGCAAAAGCCCTGGGGAGGCAAAAGTCTCTATCCCAAG	3347
QY	2442	CCCCGGGGGAGCCCTCTGTCTTTGACAAACAGCGTCTCACCTCTCCCAACCAAGTGGAGCA	2501
Db	3348	CCCCGGGGGAGCCCTCTGTCTTTGACAAACAGCGTCTCACCTCTCCCAACCAAGTGGAGCA	3407
QY	2502	GCGGAGACAGAGGCCCTCTGACAGCTGTGGTGTGGCTGGCACTGTGGGGCGTGTCTCGGAG	2561
Db	3408	GCGGAGACAGAGGCCCTCTGACAGCTGTGGTGTGGCTGGCACTGTGGGGCGTGTCTCGGAG	3467
QY	2562	CTGGGAGAGTGGCCTTGACAGAAAGCGGCGGTGGAGACTGTGGGGCTCGGCGCGAGACGAC	2621
Db	3468	CTGGGAGAGTGGCCTTGACAGAAAGCGGCGGTGGAGACTGTGGGGCTCGGCGCGAGACGAC	3527
QY	2622	GGTCCCTGCGCTGTGATGCAAGCCCATCGGCCCTGGAGACACAAAGCTGCGAGAGGCCCTG	2681
Db	3528	GGTCCCTGCGCTGTGATGCAAGCCCATCGGCCCTGGAGACACAAAGCTGCGAGAGGCCCTG	3587
QY	2682	CCCCACCTGGAGAGTGAAGCCCTGTGCACCTGCTCCAAAGACCTGGGCGCGGGATTTCA	2741
Db	3588	CCCCACCTGGAGAGTGAAGCCCTGTGCACCTGCTCCAAAGACCTGGGCGCGGGATTTCA	3647
QY	2742	GAGGCGCTCACTCAAGTGTGTGGCCACGAGAGCCGGCTGTGGGCCCGGACCAAGTGC	2801
Db	3648	GAGGCGCTCACTCAAGTGTGTGGCCACGAGAGCCGGCTGTGGGCCCGGACCAAGTGC	3707
QY	2802	CTTGACACGGAAGCCCAAGAGAGTGGACTTCTGTGCTCTGAGGCCGTGTCA	2853
Db	3708	CTTGACACGGAAGCCCAAGAGAGTGGACTTCTGTGCTCTGAGGCCGTGTCA	3759

RESULT 14
US-09-965-631-5
; Sequence 5, Application US/09965631

```

? TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
?
? FILE REFERENCE: LIX-0241-USA
?
? CURRENT APPLICATION NUMBER: US/09/965,631
?
? CURRENT FILING DATE: 2001-09-27
?
? PRIOR APPLICATION NUMBER: US 60/236,689
?
? PRIOR FILING DATE: 2000-09-29
?
? NUMBER OF SEQ. ID NOS: 7
?
? SOFTWARE: FastSeq for Windows Version 4.0
?
? SEQ ID NO 5
?
? LENGTH: 1104
?
? TYPE: DNA
?
? ORGANISM: homo sapiens
?
? OS-09-965-631-5

```

```

Query Match%: 38.2%; Score 1091; DB 36; Length 1104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCTTCTGTGGGCCTTACCTTGGCTTTGGCGGGGCAACCGCTGAGGCTCTGAG 60
|||||

```

Db	1	ATGCTTCTGCTGGGATCCTTAAACCGTGGCTTGGCGGGGAACCGGTGGAGGCTCTGAG	60
QY	61	CCAGACCGGGAGGTAGTGGTTCCTCATCCGACTGAGACCGGACATTAACGGCCGGCTAC	120
Db	61	CCAGACCGGGAGGTAGTGGTTCCTCATCCGACTGAGACCGGACATTAACGGCCGGCTAC	120
QY	121	TACTGGCGGGGTCCCGAGAGACCGGGGATGAGGAGCTCATTTTTCAGATCATAGATTT	180
Db	121	TACTGGCGGGGTCCCGAGAGACCGGGGATGAGGAGCTCATTTTTCAGATCATAGATTT	180
QY	181	CAGGAGGACTTTTACCTACACCTTGACGGCGGATGCTCAGTTCTTGGCTCCGCTTCTCC	240
Db	181	CAGGAGGACTTTTACCTACACCTTGACGGCGGATGCTCAGTTCTTGGCTCCGCTTCTCC	240
QY	241	ACTGAGCATCTGGGGGTCCCGCTCCAGGGGTCTACCGGGGGGCTTTCAGACCTGGACGC	300
Db	241	ACTGAGCATCTGGGGGTCCCGCTCCAGGGGTCTACCGGGGGGCTTTCAGACCTGGACGC	300
QY	301	TGCTTCTATTCTGGGGAGCTGAACCGGACCGAGCGGACTGTTTCTGCTGTGAGACTGTGC	360
Db	301	TGCTTCTATTCTGGGGAGCTGAACCGGACCGAGCGGACTGTTTCTGCTGTGAGACTGTGC	360
QY	361	GGGGGGGTCCCGGAGGCTTTGGGTACCGAGGGCCGAGTATGTCAATTAACCCGCTGGCC	420
Db	361	GGGGGGGTCCCGGAGGCTTTGGGTACCGAGGGCCGAGTATGTCAATTAACCCGCTGGCC	420
QY	421	TAATGCTACGGCGCGCGGCGACGACAGACAGCCAGGGCGCACACCTTCTCCAGCCCGG	480
Db	421	TAATGCTACGGCGCGCGGCGACGACAGACAGCCAGGGCGCACACCTTCTCCAGCCCGG	480
QY	481	GGTGTTCGGGGCGGGCTTCGGAGACCCACACTCTCGCTGGCGGGGTGGCTTGGGCTGG	540
Db	481	GGTGTTCGGGGCGGGCTTCGGAGACCCACACTCTCGCTGGCGGGGTGGCTTGGGCTGG	540
QY	541	AACCCCGCATCTACGGGGCCCTGGACCTTACAAAGCCGGGGCGGGGCTTCGGGGAG	600
Db	541	AACCCCGCATCTACGGGGCCCTGGACCTTACAAAGCCGGGGCGGGGCTTCGGGGAG	600
QY	601	AGTGTACCGGCGGAGTCTGGGGCGGCCAAGCGTTTCTGTCTATCCCGGGTACGTG	660
Db	601	AGTGTACCGGCGGAGTCTGGGGCGGCCAAGCGTTTCTGTCTATCCCGGGTACGTG	660
QY	661	GAGACCGGTGGTTCGGGGAGAGATCAATGTCATAGTTCAGGGCGCGGACTGGAACT	720
Db	661	GAGACCGGTGGTTCGGGGAGAGATCAATGTCATAGTTCAGGGCGCGGACTGGAACT	720
QY	721	TATCTGCTGACGCTGTGGCAACGGCGCGGACTGTACCGGCATCCAGCATCTCTCAAC	780
Db	721	TATCTGCTGACGCTGTGGCAACGGCGCGGACTGTACCGGCATCCAGCATCTCTCAAC	780
QY	781	CCCATCAACATGTTTGTGTGTCAGAGTCTGCTTTAGAGATGTGATCTTCGGGCCCAAG	840
Db	781	CCCATCAACATGTTTGTGTGTCAGAGTCTGCTTTAGAGATGTGATCTTCGGGCCCAAG	840
QY	841	GTCACCGGCATTTGGGGCTGAGCGTCCGAACTTCTGAGCCCGGAGAAAGCTGAAC	900
Db	841	GTCACCGGCATTTGGGGCTGAGCGTCCGAACTTCTGAGCCCGGAGAAAGCTGAAC	900
QY	901	AAAGTAGTGAACAAGCACCCCGAGTACTGGGACACATGCATCCTTTCACAGCGAGAC	960
Db	901	AAAGTAGTGAACAAGCACCCCGAGTACTGGGACACATGCATCCTTTCACAGCGAGAC	960
QY	961	CTGTGTGGAGCCACACTCTGAGAACCCCTGGGATGGGTGATGTGGTACATGTGTGAC	1020
Db	961	CTGTGTGGAGCCACACTCTGAGAACCCCTGGGATGGGTGATGTGGTACATGTGTGAC	1020
QY	1021	CCCAAGAAAGCTGCTGTCTATTTGAGAGAGATGGGCTTTCATCAGCTTCACACTGCC	1080
Db	1021	CCCAAGAAAGCTGCTGTCTATTTGAGAGAGATGGGCTTTCATCAGCTTCACACTGCC	1080
QY	1081	CACGAGCTGGG 1091	
Db	1081	CACGAGCTGGG 1091	

```
RESULT 15
US-09-965-631-1
; Sequence 1, Application US/09965631
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 966
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-1
```

```
Query Match      33.6%; Score 959; DB 36; Length 966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCTTCTGCTGGGCACTCCTAACCCCTGGCTTTCGCCGGGGAACCCCTGAGGCTCTGAG 60
   |||||||
DB 1 ATGCTTCTGCTGGGCACTCCTAACCCCTGGCTTTCGCCGGGGAACCCCTGAGGCTCTGAG 60

QY 61 CCAGAGCGGAGGTAGTCTGCTCCATCCAGACTGAGACCCGAGCATTAACGCGCGCGCTAC 120
   |||||||
DB 61 CCAGAGCGGAGGTAGTCTGCTCCATCCAGACTGAGACCCGAGCATTAACGCGCGCGCTAC 120

QY 121 TACTGGCGGGGTCGCCAGAGACTCCGGGGATCAGAGACTATTTTTCAGATCAGACATT 180
   |||||||
DB 121 TACTGGCGGGGTCGCCAGAGACTCCGGGGATCAGAGACTATTTTTCAGATCAGACATT 180

QY 181 CAGAGGACTTTTACCTACACACTGAGCGCGGATGCTCAGTCTTGCTCCCGGCTTCTCC 240
   |||||||
DB 181 CAGAGGACTTTTACCTACACACTGAGCGCGGATGCTCAGTCTTGCTCCCGGCTTCTCC 240

QY 241 ACTGAGCATCTGGGCGCTCCCTCCAGGGGCTCACCAGGGGCTCTTCAGACCTGCAGCG 300
   |||||||
DB 241 ACTGAGCATCTGGGCGCTCCCTCCAGGGGCTCACCAGGGGCTCTTCAGACCTGCAGCG 300

QY 301 TGCCTTATCTGGGGAGCGAGAACGCCGAGCGGACTGCTGCTGTGAGGCTGTGC 360
   |||||||
DB 301 TGCCTTATCTGGGGAGCGAGAACGCCGAGCGGACTGCTGCTGTGAGGCTGTGC 360

QY 361 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGCGCGAGTATGTATTAACCCGCTGCC 420
   |||||||
DB 361 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGCGCGAGTATGTATTAACCCGCTGCC 420

QY 421 AATGCTAGCGCGCGCGGCGCAGCGCAGACAGCCAGGGGCGACACCTTCTCCAGCGCG 480
   |||||||
DB 421 AATGCTAGCGCGCGCGGCGCAGCGCAGACAGGGGCGACACCTTCTCCAGCGCG 480

QY 481 GGTGTTCCGGGGGCGGCTTCCGAGACCCACCTCTGCTGCGGGGTGCGGGCTGG 540
   |||||||
DB 481 GGTGTTCCGGGGGCGGCTTCCGAGACCCACCTCTGCTGCGGGGTGCGGGCTGG 540

QY 541 AACCCGCGCATCTACAGGGGCTGAGACCTTAAGCGCGCGGCGGCGGCTTCGGGGAG 600
   |||||||
DB 541 AACCCGCGCATCTACAGGGGCTGAGACCTTAAGCGCGCGGCGGCGGCTTCGGGGAG 600

QY 601 AGTGTAGCGCGCGCAGGCTTGGGGCGCGCAAGCGTTTGTGTATCCCGGCTACGTG 660
   |||||||
DB 601 AGTGTAGCGCGCGCAGGCTTGGGGCGCGCAAGCGTTTGTGTATCCCGGCTACGTG 660

QY 661 GAGACGCTGGTGTGCGGGGAGAGTCAATAGTTCACGGCGGCGGACTGGAACAT 720
   |||||||
DB 661 GAGACGCTGGTGTGCGGGGAGAGTCAATAGTTCACGGCGGCGGACTGGAACAT 720
```

```
QY 721 TATCGTGAAGCTGTGAGCAAGCGGCGGAGACTTACCCCATCCAGACATCTCAAC 780
   |||||||
DB 721 TATCGTGAAGCTGTGAGCAAGCGGCGGAGACTTACCCCATCCAGACATCTCAAC 780

QY 781 CCCATCAACATCGTTGTGAGTCAAGGTGCTGCTTCTTAAGATGCTGACTCCGGGCCAAG 840
   |||||||
DB 781 CCCATCAACATCGTTGTGAGTCAAGGTGCTGCTTCTTAAGATGCTGACTCCGGGCCAAG 840

QY 841 GTCACCGGCAATGGGGCCCTTACGCTGCGCACTTCTGTGCTGGGCAAGAGAAGCTGAAC 900
   |||||||
DB 841 GTCACCGGCAATGGGGCCCTTACGCTGCGCACTTCTGTGCTGGGCAAGAGAAGCTGAAC 900

QY 901 AAAGTGAAGTGAACAGCACCCCGAGTACAGGACACTGCACTCTCTTCACAGGACAGA 959
   |||||||
DB 901 AAAGTGAAGTGAACAGCACCCCGAGTACAGGACACTGCACTCTCTTCACAGGACAGA 959
```

```
Search completed: May 16, 2003, 05:38:15
Job time : 5618 secs
```

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 03:53:22 ; Search time 12839 Seconds
(without alignments)
1149.891 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgcttcgtgctgagcctcct.....gctcctgagcgcgtgctga 2853

Scoring table: OLIGO NWC
Gapop 60.0 , Gapext 60.0

Searched: 6404235 seqs, 2587356060 residues

Word size : 24

Total number of hits satisfying chosen parameters: 24

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2802	98.2	2853	8 US-10-391-364-76	Sequence 76, Appl
2	2802	98.2	2853	8 US-10-391-364-78	Sequence 78, Appl
3	2802	98.2	2853	8 US-10-170-235-10538	Sequence 10538, A
4	2802	98.2	2853	10 US-60-466-411-7	Sequence 7, Appl
5	2469	86.5	2853	11 US-60-453-135-7278	Sequence 7278, Ap
6	2469	86.5	2853	11 US-60-453-050-7278	Sequence 32, Appl
7	2196	77.0	2930	9 US-10-311-035-32	Sequence 77890, A
8	150	5.3	201	11 US-60-453-135-77889	Sequence 77890, A
9	150	5.3	201	11 US-60-453-135-77891	Sequence 77891, A
10	150	5.3	201	11 US-60-453-135-77892	Sequence 77892, A
11	150	5.3	201	11 US-60-453-135-77893	Sequence 77893, A
12	150	5.3	201	11 US-60-453-135-77894	Sequence 77894, A
13	150	5.3	201	11 US-60-453-050-77889	Sequence 77889, A
14	150	5.3	201	11 US-60-453-050-77890	Sequence 77890, A
15	150	5.3	201	11 US-60-453-050-77891	Sequence 77891, A
16	150	5.3	201	11 US-60-453-050-77892	Sequence 77892, A
17	150	5.3	201	11 US-60-453-050-77893	Sequence 77893, A
18	127	4.5	178	11 US-60-453-135-77892	Sequence 77892, A
19	127	4.5	178	11 US-60-453-050-77892	Sequence 77892, A
20	126	4.4	177	11 US-60-453-135-77895	Sequence 77895, A
21	126	4.4	177	11 US-60-453-050-77895	Sequence 77895, A
22	100	3.5	115	11 US-60-453-135-77896	Sequence 77896, A

23 100 3.5 115 11 US-60-453-050-77896 Sequence 77896, A
24 74 2.6 5714 9 US-10-144-771-12354 Sequence 12354, A

ALIGNMENTS

```
RESULT 1
US-10-391-364-76
; Sequence 76, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352.
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 55577 MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MP103-0150MINI
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2853)
; US-10-391-364-76
Query Match 98.2% Score 2802; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 ATGCTTCGCTGGGAGCTTACCTGAGCTTCCCGGGGAGAACCGCTGAGGCTTGAG 60
DB 1 AMGCTTCGCTGGGAGCTTACCTGAGCTTCCCGGGGAGAACCGCTGAGGCTTGAG 60
QY 61 CCAAGGCGGAGGTAGTGTTCCTCCATCCGACTGACCCCGGACATTAAAGGCGCGCTAC 120
DB 61 CCAAGGCGGAGGTAGTGTTCCTCCATCCGACTGACCCCGGACATTAAAGGCGCGCTAC 120
QY 121 TACGCGCGGCTCCCGAGGACTCCGGGAGTACGAGACTCAATTTTTCAGATCAGACATT 180
DB 121 TACGCGCGGCTCCCGAGGACTCCGGGAGTACGAGACTCAATTTTTCAGATCAGACATT 180
QY 181 CAGAGGACTTTTACCTACACCTGACCCCGAGTCTCAGTTCTTGCTCCGCTTCTCC 240

|||||
Db 181 CAGAGGACTTTTACCTACACCTGAGCCGGATGCTCAGTTCTTGCTCCGCTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGCTCTTAGACCTCCGACGC 300
Db 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGCTCTTAGACCTCCGACGC 300
QY 301 TGTCTATTCTGGGAGAGTGAACGGCCGAGCTGTTGGCTGTGAGACTGTG 360
Db 301 TGTCTATTCTGGGAGAGTGAACGGCCGAGCTGTTGGCTGTGAGACTGTG 360
QY 361 GGGGGGCTCCGGAGGCTTTGGCTACCGAGGGCGGAGTATGTCATTAGCCGCTGCC 420
Db 361 GGGGGGCTCCGGAGGCTTTGGCTACCGAGGGCGGAGTATGTCATTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGGGGAGCAGCAGACGGGGGCGACACCTTCCTCAAGCGCGG 480
Db 421 AATGCTAGCGCGCGGGGGGAGCAGCAGACGGGGGCGACACCTTCCTCAAGCGCGG 480
QY 481 GGTGTTCCGGGCGGGCTTCGAGAGACCCACCTCTGCTGCGGGGTGGCTGGGCTGG 540
Db 481 GGTGTTCCGGGCGGGCTTCGAGAGACCCACCTCTGCTGCGGGGTGGCTGGGCTGG 540
QY 541 AATGCTAGCGCGCGGGGGGAGCAGCAGACGGGGGCGACACCTTCCTCAAGCGCGG 600
Db 541 AATGCTAGCGCGCGGGGGGAGCAGCAGACGGGGGCGACACCTTCCTCAAGCGCGG 600
QY 601 AGTGTAGCGCGCGGAGGCTGCGGGCGGCAAGGTTTCTGTCATCCGCGGTACG 660
Db 601 AGTGTAGCGCGCGGAGGCTGCGGGCGGCAAGGTTTCTGTCATCCGCGGTACG 660
QY 661 GAGAGCGTGTGTGCGCGGAGCAATGTCATGTCAGGCGGCGACTGGAAAT 720
Db 661 GAGAGCGTGTGTGCGCGGAGCAATGTCATGTCAGGCGGCGACTGGAAAT 720
QY 721 TATGCTGAGCGGCTGGGCAAGGGGGGCGAGCTACCGGCACTCCACATCTC 780
Db 721 TATGCTGAGCGGCTGGGCAAGGGGGGCGAGCTACCGGCACTCCACATCTC 780
QY 781 CCCATCAACATCTGTTGTGTCAGAGTGTCTGTTAGAGATGTCGAGCTCCGGGCCAAG 840
Db 781 CCCATCAACATCTGTTGTGTCAGAGTGTCTGTTAGAGATGTCGAGCTCCGGGCCAAG 840
QY 841 GTACACGGCAATGCGGCGCTGAGCGCTGCGCACTTCTGTGCTGCGGAGAAAGCTG 900
Db 841 GTACACGGCAATGCGGCGCTGAGCGCTGCGCACTTCTGTGCTGCGGAGAAAGCTG 900
QY 901 AAGGTAGTGAACAGCAACCCGAGTACTGGAGACTGCACTCTTCTACACAGCGAGAC 960
Db 901 AAGGTAGTGAACAGCAACCCGAGTACTGGAGACTGCACTCTTCTACACAGCGAGAC 960
QY 961 CTGTGTGAGGCGACACCTGTGACACCTGGGCAATGCTATGCTGAGTACATGTG 1020
Db 961 CTGTGTGAGGCGACACCTGTGACACCTGGGCAATGCTATGCTGAGTACATGTG 1020
QY 1021 CCCAAGAGAAAGTCTCTGTATTTAGAGAGATGGCTTCATCAGCTTCCACACTGCC 1080
Db 1021 CCCAAGAGAAAGTCTCTGTATTTAGAGAGATGGCTTCATCAGCTTCCACACTGCC 1080
QY 1081 CACGAGTGGGGCGAGTGTTCACATGCCCCATGACATGGAAGCTGTGAGAGGTG 1140
Db 1081 CACGAGTGGGGCGAGTGTTCACATGCCCCATGACATGGAAGCTGTGAGAGGTG 1140
QY 1141 TTTGGGAGAGCTCCGAGCCAAACCATGATGTCCCGAACCTTCATCCAGATGCACTGCC 1200
Db 1141 TTTGGGAGAGCTCCGAGCCAAACCATGATGTCCCGAACCTTCATCCAGATGCACTGCC 1200
QY 1201 AATGCTAGCGCGCGGGGGGAGCAGCAGACGGGGGCGACACCTTCCTCAAGCGCGG 1260
Db 1201 AATGCTAGCGCGCGGGGGGAGCAGCAGACGGGGGCGACACCTTCCTCAAGCGCGG 1260
QY 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCCCTGCGCAAGATGTCGCGGGGCGC 1320
Db 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCCCTGCGCAAGATGTCGCGGGGCGC 1320

Db 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCCCTGCGCAAGATGTCGCGGGGCGC 1320
QY 1321 AGCTACACCTGAGGAGAGAGGAGGCTGTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
Db 1321 AGCTACACCTGAGGAGAGAGGAGGCTGTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
QY 1381 TACATGCACTACTGACCAAGCTGTGTGTCACCGGGAAAGGCCAAGAGATGATGTG 1440
Db 1381 TACATGCACTACTGACCAAGCTGTGTGTCACCGGGAAAGGCCAAGAGATGATGTG 1440
QY 1441 CAGACCCGCACTTCCCTGGGCGCATGGGCAACACCTGTGGCGAGAGCACTCTGCTC 1500
Db 1441 CAGACCCGCACTTCCCTGGGCGCATGGGCAACACCTGTGGCGAGAGCACTCTGCTC 1500
QY 1501 AAGGGGCGCTGCGTGAAGAGACACAACTCAACAGCAGAGGTGATGTTCTGCGGC 1560
Db 1501 AAGGGGCGCTGCGTGAAGAGACACAACTCAACAGCAGAGGTGATGTTCTGCGGC 1560
QY 1561 AATGGGATCCCTATGAGCCCTGTCGCGCACATGTGTTGGGGGCTGACCTGCGCAGG 1620
Db 1561 AATGGGATCCCTATGAGCCCTGTCGCGCACATGTGTTGGGGGCTGACCTGCGCAGG 1620
QY 1621 AGGAGTGTGACCAACCCCAACCCCTGCAAGGGGGGCAAGTACTGGAGAGTGAAGGTG 1680
Db 1621 AGGAGTGTGACCAACCCCAACCCCTGCAAGGGGGGCAAGTACTGGAGAGTGAAGGTG 1680
QY 1681 AATACCGATCTCTGCAATCTGAGAGCCCTGCGCACCTCAGCTCCGAAAGAGCTTCGG 1740
Db 1681 AATACCGATCTCTGCAATCTGAGAGCCCTGCGCACCTCAGCTCCGAAAGAGCTTCGG 1740
QY 1741 GAGGAGCACTGTGAGGCTTTCAAGGGCTTCAAGCAACACAGCAGCCAGCTCACTTGGCC 1800
Db 1741 GAGGAGCACTGTGAGGCTTTCAAGGGCTTCAAGCAACACAGCAGCCAGCTCACTTGGCC 1800
QY 1801 GTGGAGTGGTGGCCCAAGTACTCCGGGCTGTCCCGGGAGCAAGTGAAGCTATCTG 1860
Db 1801 GTGGAGTGGTGGCCCAAGTACTCCGGGCTGTCCCGGGAGCAAGTGAAGCTATCTG 1860
QY 1861 CGAGCCATGAGCACTGCTACTTCTATGTGTCGAGCAACCAAGGTGTGAGCGCACGCTG 1920
Db 1861 CGAGCCATGAGCACTGCTACTTCTATGTGTCGAGCAACCAAGGTGTGAGCGCACGCTG 1920
QY 1921 TGTCTCTGACTCCACCTCGCTGTGTGTCGAAGGCAAGTGTATCAAGGCTGTGAT 1980
Db 1921 TGTCTCTGACTCCACCTCGCTGTGTGTCGAAGGCAAGTGTATCAAGGCTGTGAT 1980
QY 1981 GGAACCTGGGCTCCAAAGAGAGATTTGCAAGTGTGGGCTGTGGGGAGACAATAG 2040
Db 1981 GGAACCTGGGCTCCAAAGAGAGATTTGCAAGTGTGGGCTGTGGGGAGACAATAG 2040
QY 2041 AGCTGCAAGAGGTGACTGACTCTTCAACCAAGCCATGATGCTACATTTCTGTG 2100
Db 2041 AGCTGCAAGAGGTGACTGACTCTTCAACCAAGCCATGATGCTACATTTCTGTG 2100
QY 2101 GGCATCCCGGAGGGGCTTCAAGCATGCAATCCCGGAGCGGTTCAAAGGCTGTATC 2160
Db 2101 GGCATCCCGGAGGGGCTTCAAGCATGCAATCCCGGAGCGGTTCAAAGGCTGTATC 2160
QY 2161 GGGGATGACAATCTACTGCTCTGAAGAAACAGCAGGCAAGTACTCTCAAGCGGAT 2220
Db 2161 GGGGATGACAATCTACTGCTCTGAAGAAACAGCAGGCAAGTACTCTCAAGCGGAT 2220
QY 2221 TTTGTTGTGTGGGCGGTGAGCGGAGCTGTGTGAAGGGGAGTCTGCTCGGTACGC 2280
Db 2221 TTTGTTGTGTGGGCGGTGAGCGGAGCTGTGTGAAGGGGAGTCTGCTCGGTACGC 2280
QY 2281 GGCAGGGGAGACGGGTGAGAGCTGTGAGAGCTTCCGCGCATCTGAGAGCGCTGAC 2340
Db 2281 GGCAGGGGAGACGGGTGAGAGCTGTGAGAGCTTCCGCGCATCTGAGAGCGCTGAC 2340
QY 2341 GTGAGAGTCTCTCGTGGGAGAGTGAACAGCCCGGGGTCCGCTACTCTTATCTG 2400
Db 2341 GTGAGAGTCTCTCGTGGGAGAGTGAACAGCCCGGGGTCCGCTACTCTTATCTG 2400

```

QY 2401 CCCAAGAGCCTCGGAGGAGCAAGTCTCATCCCAAGAGACCCCGGGACCCCTCTGTC 2460
      |||||||
Db 2401 CCCAAGAGCCTCGGAGGAGCAAGTCTCATCCCAAGAGACCCCGGGACCCCTCTGTC 2460
QY 2461 TTGGACAAAGCGGTCCTCACCCTCTCCAAACAGAGTGAGAGCCGAGACAGAGCCCT 2520
      |||||||
Db 2461 TTGGACAAAGCGGTCCTCACCCTCTCCAAACAGAGTGAGAGCCGAGACAGAGCCCT 2520
QY 2521 GCAACGCTGGGTGCTGGACAGTGGGGGCGCTGCTCCGAGCTGGCGAGTGGGCTGAC 2580
      |||||||
Db 2521 GCAACGCTGGGTGCTGGACAGTGGGGGCGCTGCTCCGAGCTGGCGAGTGGGCTGAC 2580
QY 2581 AAGCGCGGCTGAGCTGTGGGGCTCCCGCGGACGACAGCGTCCCTGCTGTGATGCA 2640
      |||||||
Db 2581 AAGCGCGGCTGAGCTGTGGGGCTCCCGCGGACGACAGCGTCCCTGCTGTGATGCA 2640
QY 2641 GCCCATCGGCGGCTGGAGACAAAGCTCGGGGAGACCCCTGCCCCACCTGGAGCTCAGC 2700
      |||||||
Db 2641 GCCCATCGGCGGCTGGAGACAAAGCTCGGGGAGACCCCTGCCCCACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGCTCAAGAGCTGCGGCGGGGATTTCAAGAGCGCTCACTCAAGTGT 2760
      |||||||
Db 2701 GCCTGGTCAACCTGCTCAAGAGCTGCGGCGGGGATTTCAAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGCAGGAGCGCGGCTGCTGCGCGGAGCAAGTGCACCTTGACCCGCAAGCCCGCAG 2820
      |||||||
Db 2761 GTGGGCGCAGGAGCGCGGCTGCTGCGCGGAGCAAGTGCACCTTGACCCGCAAGCCCGCAG 2820
QY 2821 GAGCTGAGCTTGTGCTGCTGCGGAGCGGCTGCTCA 2853
      |||||||
Db 2821 GAGCTGAGCTTGTGCTGCTGCGGAGCGGCTGCTCA 2853

```

```

RESULT 2
US-10-391-364-78
; Sequence 78, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 1658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; FILE REFERENCE: MPI03-0190MNTM
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2853)
US-10-391-364-78

Query Match          98.2%; Score 2802; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTTCGTGGGATCTTAACCTGCTTTCGCGGGGCGCAACCGCTGGAGGCTGTGAG 60
      |||||||
Db 1 ATGCTTTCGTGGGATCTTAACCTGCTTTCGCGGGGCGCAACCGCTGGAGGCTGTGAG 60
QY 61 CCAGAGCGGAGGTAGTGTCCATCCGACTGGACCCGAGACATTAACGGCGCGCTAC 120
      |||||||
Db 61 CCAGAGCGGAGGTAGTGTCCATCCGACTGGACCCGAGACATTAACGGCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGAGCTCCGGGGATCAGGACTATTTTCAGATCAGCATTT 180
      |||||||
Db 121 TACTGGCGGGTCCCGAGAGCTCCGGGGATCAGGACTATTTTCAGATCAGCATTT 180
QY 181 CAGAGGACTTTTACCTACCTGACGCGGGATGCTAGTTCTTGCTCCGCTTCTCC 240
      |||||||
Db 181 CAGAGGACTTTTACCTACCTGACGCGGGATGCTAGTTCTTGCTCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGGGTCCTCCCTCAGAGGCTCACCGGGGCTCTTCAGACTCGACGC 300
      |||||||
Db 241 ACTGAGCATCTGGGGGTCCTCCCTCAGAGGCTCACCGGGGCTCTTCAGACTCGACGC 300
QY 301 TGCTTCTATTTCTGGGAGCTGAACGCGGAGCTGCTTGTGAGCTGTGC 360
      |||||||
Db 301 TGCTTCTATTTCTGGGAGCTGAACGCGGAGCTGCTTGTGAGCTGTGC 360
QY 361 GGGGGGCTCCGCGAGCTTTGGCTACCGAGGCGGAGATGTATGTATAGCCGCTGCC 420
      |||||||
Db 361 GGGGGGCTCCGCGAGCTTTGGCTACCGAGGCGGAGATGTATGTATAGCCGCTGCC 420
QY 421 AATGCTACGCGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGCGG 480
      |||||||
Db 421 AATGCTACGCGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGCGG 480
QY 481 GGTGTTCGCGGCGGCTTCCGAGAGCCCACTCTGCTGCGGGGTGGGCTGCG 540
      |||||||
Db 481 GGTGTTCGCGGCGGCTTCCGAGAGCCCACTCTGCTGCGGGGTGGGCTGCG 540
QY 541 AACCCGCGATCTTACGAGGCGCTTGAGCCCTTACAAAGCGGCGGCGGCTTCCG 600
      |||||||
Db 541 AACCCGCGATCTTACGAGGCGCTTGAGCCCTTACAAAGCGGCGGCGGCTTCCG 600
QY 601 AGTGTAGCGGCGGAGGCTGAGGCGGCGGAGCGGAGCGGCTTCTGCTATCCCG 660
      |||||||
Db 601 AGTGTAGCGGCGGAGGCTGAGGCGGCGGAGCGGAGCGGCTTCTGCTATCCCG 660
QY 661 GAGAGCGTGTGTGCGGAGGAGGAGTCAATGTAAGTTTCAAGGCGGCGGAGGAG 720
      |||||||
Db 661 GAGAGCGTGTGTGCGGAGGAGGAGTCAATGTAAGTTTCAAGGCGGCGGAGGAG 720
QY 721 TATCTGTGAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
      |||||||
Db 721 TATCTGTGAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 CCATCAACATCGTTGTGTCAAGGTGCTTCTTGAAGATGTGACTCCGCGGCGCAAG 840
      |||||||
Db 781 CCATCAACATCGTTGTGTCAAGGTGCTTCTTGAAGATGTGACTCCGCGGCGCAAG 840
QY 841 GTACCGGCAATGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
      |||||||
Db 841 GTACCGGCAATGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900

```


LENGTH: 2853
TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-10538

Query Match 98.2% Score 2802; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCTTCTGCTGGGATCTTACCTACCCGCTTTGGCCGGGGAACCGCTGAGGCTTGAG 60
    |||||||
Db 1 ATGCTTCTGCTGGGATCTTACCTACCCGCTTTGGCCGGGGAACCGCTGAGGCTTGAG 60
QY 61 CCAGAGCGGAGGTAGTGTCTTCCATCCGACTGAGCCGCGCATTTAACGCGCGCTTAC 120
    |||||||
Db 61 CCAGAGCGGAGGTAGTGTCTTCCATCCGACTGAGCCGCGCATTTAACGCGCGCTTAC 120
QY 121 TACTGGCGGGGTCGCCGAGGACTCCGGGGATCAGGGACTATTTTCAGATCACAGCATTT 180
    |||||||
Db 121 TACTGGCGGGGTCGCCGAGGACTCCGGGGATCAGGGACTATTTTCAGATCACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGAGCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
    |||||||
Db 181 CAGGAGGACTTTTACCTACACCTGAGCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCCTCCAGGGGCTCAGCGGGGCTTCCAGACCTGCGAGCG 300
    |||||||
Db 241 ACTGAGCATCTGGGCGTCCCTCCCTCCAGGGGCTCAGCGGGGCTTCCAGACCTGCGAGCG 300
QY 301 TCGTTTATTTCTGGGAGCTGAAAGCCGAGCCGAGCTGTTGCTGCTGTGAGCCCTGTGC 360
    |||||||
Db 301 TCGTTTATTTCTGGGAGCTGAAAGCCGAGCCGAGCTGTTGCTGCTGTGAGCCCTGTGC 360
QY 361 GGGGGGCTCCGCGGAGCTTTGGCTACCGAGGCGCGAGTGTCTTACGCGCGCTGCC 420
    |||||||
Db 361 GGGGGGCTCCGCGGAGCTTTGGCTACCGAGGCGCGAGTGTCTTACGCGCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCCAGGGGCGACACCTTCTCCAGCGCGCG 480
    |||||||
Db 421 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCCAGGGGCGACACCTTCTCCAGCGCGCG 480
QY 481 GGTGTTCCGGGGCGGCTTCCGAGAGCCCGACCTCTGCGGGGTGCGCTGCGCGCTGCG 540
    |||||||
Db 481 GGTGTTCCGGGGCGGCTTCCGAGAGCCCGACCTCTGCGGGGTGCGCTGCGCGCTGCG 540
QY 541 AACCCCGCATCTTACGCGGCGCTGAGCCCTTACAAAGCCGCGGGCGGCGCTTCCGGGAG 600
    |||||||
Db 541 AACCCCGCATCTTACGCGGCGCTGAGCCCTTACAAAGCCGCGGGCGGCGCTTCCGGGAG 600
QY 601 AGTCGTAAGCCGCGAGGTCTGCGGCGCGCAAGCGTTTGTCTATCCCGGGAAGCTG 660
    |||||||
Db 601 AGTCGTAAGCCGCGAGGTCTGCGGCGCGCAAGCGTTTGTCTATCCCGGGAAGCTG 660
QY 661 GAGAGCTGTGGTGGCGGAGAGTCAATGTCAAGTTTCCAGAGGCGCGGACCTTGAACAT 720
    |||||||
Db 661 GAGAGCTGTGGTGGCGGAGAGTCAATGTCAAGTTTCCAGAGGCGCGGACCTTGAACAT 720
QY 721 TATCTGCTAGCGCTGTGCGCAAGGCGCGGAGCTTACCGGCATCCAGACTCTCTCAAC 780
    |||||||
Db 721 TATCTGCTAGCGCTGTGCGCAAGGCGCGGAGCTTACCGGCATCCAGACTCTCTCAAC 780
QY 781 CCCATCAACATGTGTGTGTCAGAGTGTCTTTCAGATGTGACTCCGGGCCCAAG 840
    |||||||
Db 781 CCCATCAACATGTGTGTGTCAGAGTGTCTTTCAGATGTGACTCCGGGCCCAAG 840
QY 841 GTACACGCGCATGTGCGGCGCTGAGCGTGGCGCACTTCTGCTGGCAGAGAAGTGTAC 900
    |||||||
Db 841 GTACACGCGCATGTGCGGCGCTGAGCGTGGCGCACTTCTGCTGGCAGAGAAGTGTAC 900
QY 901 AAAGTGAAGTGAAGACACCCGAGTACTGGGACACTGCGCATCTTTCACAGGAGGAGAC 960
    |||||||
Db 901 AAAGTGAAGTGAAGACACCCGAGTACTGGGACACTGCGCATCTTTCACAGGAGGAGAC 960
```

```
QY 961 CTGTGTGAGGACCAACCCGTGTGACACCCGTGGGACGTGATGTGGTACCATGTGTAC 1020
    |||||||
Db 961 CTGTGTGAGGACCAACCCGTGTGACACCCGTGGGACGTGATGTGGTACCATGTGTAC 1020
QY 1021 CCCAAGAGAGGTGCTGTGCTGCTATGAGAGCAGTGGGCTTCCATCGACCTTCAACCTGCC 1080
    |||||||
Db 1021 CCCAAGAGAGGTGCTGTGCTGCTATGAGAGCAGTGGGCTTCCATCGACCTTCAACCTGCC 1080
QY 1081 CACGAGCTGGGCGCACGTGTCAACATGCCCCATGACAAATGTGAAAGTGTGAGAGGTG 1140
    |||||||
Db 1081 CACGAGCTGGGCGCACGTGTCAACATGCCCCATGACAAATGTGAAAGTGTGAGAGGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCCCAACCATGATGTGCCGACCTTATCCATGTGACCGTGGC 1200
    |||||||
Db 1141 TTTGGGAAGCTCCGAGCCCAACCATGATGTGCCGACCTTATCCATGTGACCGTGGC 1200
QY 1201 AACCCCTGTGAGCTGTGAGTGTCTGCTATACACCGACTTCTGAGAGGCGGAGCGST 1260
    |||||||
Db 1201 AACCCCTGTGAGCTGTGAGTGTCTGCTATACACCGACTTCTGAGAGGCGGAGCGST 1260
QY 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTCCGAGAGATCTCCGGCGGCTC 1320
    |||||||
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTCCGAGAGATCTCCGGCGGCTC 1320
QY 1321 AGCTACACCCGTGAGCGACGAGTGGAGCTTTGGCGTGGGCTTCCAAAGCTGTCT 1380
    |||||||
Db 1321 AGCTACACCCGTGAGCGACGAGTGGAGCTTTGGCGTGGGCTTCCAAAGCTGTCT 1380
QY 1381 TACATGACGTACTGACCAACGCTGTGTGACACCGGAGGCAAGGCGCAGATGCTGTGC 1440
    |||||||
Db 1381 TACATGACGTACTGACCAACGCTGTGTGACACCGGAGGCAAGGCGCAGATGCTGTGC 1440
QY 1441 CAGACCCGCGCACTTCCCTGGGCGGATGTGCACAGCTGTGCGAGAGGCGACCTGTGCTC 1500
    |||||||
Db 1441 CAGACCCGCGCACTTCCCTGGGCGGATGTGCACAGCTGTGCGAGAGGCGACCTGTGCTC 1500
QY 1501 AAAGGGGCTCGTGGAGAGACACAACTCTCAAGAGCAGAGGAGTGGTGTCTGGGCGC 1560
    |||||||
Db 1501 AAAGGGGCTCGTGGAGAGACACAACTCTCAAGAGCAGAGGAGTGGTGTCTGGGCGC 1560
QY 1561 AAATGGATCCCTATGAGCCCTGTCTCGGCGACATGTGTGGGGCGTGCAGCTGGCCAG 1620
    |||||||
Db 1561 AAATGGATCCCTATGAGCCCTGTCTCGGCGACATGTGTGGGGCGTGCAGCTGGCCAG 1620
QY 1621 AGGCGTGCACCAACCCCGACCCCTGCCAGCGGGGCGAAGTCTCGAGAGGAGTGAAGGTG 1680
    |||||||
Db 1621 AGGCGTGCACCAACCCCGACCCCTGCCAGCGGGGCGAAGTCTCGAGAGGAGTGAAGGTG 1680
QY 1681 AAATACCGATCTGTCAATCTGAGACCCCTGCCAGCTCAGGCTCCGGAAGAGGCTTCC 1740
    |||||||
Db 1681 AAATACCGATCTGTCAATCTGAGACCCCTGCCAGCTCAGGCTCCGGAAGAGGCTTCC 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCACGCGCTACAAACAGCAACCAACGCGCTACTCTCGCC 1800
    |||||||
Db 1741 GAGGAGCAGTGTGAGGCTTTCACGCGCTACAAACAGCAACCAACGCGCTACTCTCGCC 1800
QY 1801 GTGGATGGTGGCCCAAGTACTCGGGGTGTCTCCCGGAGCAAGTGCAGTCAATCTGC 1860
    |||||||
Db 1801 GTGGATGGTGGCCCAAGTACTCGGGGTGTCTCCCGGAGCAAGTGCAGTCAATCTGC 1860
QY 1861 CGAGCCATGTGCACTGTGCTACTTCTATGTGTGCGACCCCAAGGTGTGAGAGCGCACG 1920
    |||||||
Db 1861 CGAGCCATGTGCACTGTGCTACTTCTATGTGTGCGACCCCAAGGTGTGAGAGCGCACG 1920
QY 1921 TGTCTCTCTGACTCCACTCTGCTGTGTGTCCAAAGGCAAGTGTCAAGGCTGTGCTGTGAT 1980
    |||||||
Db 1921 TGTCTCTCTGACTCCACTCTGCTGTGTGTCCAAAGGCAAGTGTCAAGGCTGTGCTGTGAT 1980
QY 1981 GGGAACTGGGCTCCAAAGAAAGATTCGACAAAGTGTGGGTGTGGGGGAGGACATTAAG 2040
    |||||||
Db 1981 GGGAACTGGGCTCCAAAGAAAGATTCGACAAAGTGTGGGTGTGGGGGAGGACATTAAG 2040
QY 2041 AGCTGCAAGAAAGGTGACTGACCTCTTTCACCAAGCCCATGCAATTTTCGTGTG 2100
    |||||||
```


|||||
Db 961 CTGTGTGGAGCCACACCTGTGACACCCCTGGGCACTGATGTGGTACATGTGTGAC 1020
OY 1021 CCCAAGAGAAAGCTCTCTGTATGTAGAGAGATGGGCTTCCATCAGCTTACCACCTGCC 1080
Db 1021 CCCAAGAGAAAGCTCTCTGTATGTAGAGAGATGGGCTTCCATCAGCTTACCACCTGCC 1080
OY 1081 CACGAGCTGGGGCCACGTGTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1140
Db 1081 CACGAGCTGGGGCCACGTGTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1140
OY 1141 TTGTGGAGAGCTCCGAGCCAAACCAATGATGTCCCGACCTCAATCCAGATGACCGTGGC 1200
Db 1141 TTGTGGAGAGCTCCGAGCCAAACCAATGATGTCCCGACCTCAATCCAGATGACCGTGGC 1200
OY 1201 AACCCCTGTGAGCTGCTGATGCTGATGATGACGACTTCTGTGACAGACGGGACGGT 1260
Db 1201 AACCCCTGTGAGCTGCTGATGCTGATGATGACGACTTCTGTGACAGACGGGACGGT 1260
OY 1261 GACTGCTCTGTGAGCAACCCAGCAAGCCCATCTCCCTGCGGAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTGTGAGCAACCCAGCAAGCCCATCTCCCTGCGGAGGATCTGCCGGGCGCC 1320
OY 1321 AGCTAACACCTGAGCCAGCAGTGTGAGTGTGCTTTTGGGTGGCTTCCAAAGCCCTGTCT 1380
Db 1321 AGCTAACACCTGAGCCAGCAGTGTGAGTGTGCTTTTGGGTGGCTTCCAAAGCCCTGTCT 1380
OY 1381 TACATGAGTACTGACCAACAGTGTGTGACCGGAGCAAGGAGCAAGATGATGTGTC 1440
Db 1381 TACATGAGTACTGACCAACAGTGTGTGACCGGAGCAAGGAGCAAGATGATGTGTC 1440
OY 1441 CAGACCCGACCTTCCCTGTGGGCGATGTGACACAGCTGTGGCGAGGCAAGCTCTGCTC 1500
Db 1441 CAGACCCGACCTTCCCTGTGGGCGATGTGACACAGCTGTGGCGAGGCAAGCTCTGCTC 1500
OY 1501 AAAGGGGCGTGGTGGAGAGACACACTTCAACACAGAGGTGATGCTTCTGTGGGCG 1560
Db 1501 AAAGGGGCGTGGTGGAGAGACACACTTCAACACAGAGGTGATGCTTCTGTGGGCG 1560
OY 1561 AAATGGGATCCCTATGTGGGCGCTGTGCGACATGTGTGGGGGCGTGTGACGCTGGGCGAG 1620
Db 1561 AAATGGGATCCCTATGTGGGCGCTGTGCGACATGTGTGGGGGCGTGTGACGCTGGGCGAG 1620
OY 1621 AGCAGTGTGACCAACCCACCCCTGTGCAAGGGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
Db 1621 AGCAGTGTGACCAACCCACCCCTGTGCAAGGGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
OY 1681 AAATACGATCCGCAATGTGAGCGCTGCGCCAGCTCAGGCTCCGGAAGAGCTTCCGG 1740
Db 1681 AAATACGATCCGCAATGTGAGCGCTGCGCCAGCTCAGGCTCCGGAAGAGCTTCCGG 1740
OY 1741 GAGGAGCAGTGTGAGGCTTTCACAGGCTTACACACAGACAGACCGGCTCACTCTCGCC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCACAGGCTTACACACAGACAGACCGGCTCACTCTCGCC 1800
OY 1801 GTGGAGTGGGTGCCCAAGTACTCCGGGTCTCCCGGGGCAAGTGAAGCTCATCTGTC 1860
Db 1801 GTGGAGTGGGTGCCCAAGTACTCCGGGTCTCCCGGGGCAAGTGAAGCTCATCTGTC 1860
OY 1861 CGAGCCAAATGGGCACTGCTTATGTGTGAGCAACCCAAAGTGTGTGAACGGGACGCTG 1920
Db 1861 CGAGCCAAATGGGCACTGCTTATGTGTGAGCAACCCAAAGTGTGTGAACGGGACGCTG 1920
OY 1921 TGTCTCTGACTGCAACCTCCTGTGTGTCAAGGCAAGTGTCAAAAGGCTGGTGTGAT 1980
Db 1921 TGTCTCTGACTGCAACCTCCTGTGTGTCAAGGCAAGTGTCAAAAGGCTGGTGTGAT 1980
OY 1981 GGAAGCTGGGCTCCAAAGAGAGATTCGACAAAGTGTGGGGGTGTGGGGGGAGACAATAAG 2040
Db 1981 GGAAGCTGGGCTCCAAAGAGAGATTCGACAAAGTGTGGGGGTGTGGGGGGAGACAATAAG 2040
OY 2041 AGCTGCAAGAGGTGACTGTGCACTTTCACCAAGCCCATGCAATGCTTCAATTTGCTGTG 2100
|||||

Db 2041 AGCTGCAAGAGGTGACTGTGCACTTTCACCAAGCCCATGCAATGCTTCAATTTCTGTG 2100
OY 2101 GCCATCCCCGAGAGGCGCTCAAGCATGCAATCGCCAGCGGGTTACAAAGGCTGTATC 2160
Db 2101 GCCATCCCCGAGAGGCGCTCAAGCATGCAATCGCCAGCGGGTTACAAAGGCTGTATC 2160
OY 2161 GGGGATGACACTTACTGCTGTGAAGAACAGCCAAAGGCAAGTACTCTCTCAAGCGGAT 2220
Db 2161 GGGGATGACACTTACTGCTGTGAAGAACAGCCAAAGGCAAGTACTCTCTCAAGCGGAT 2220
OY 2221 TTCGTGTGTGCGCGGTGAGCGGAGCTGTGTGTGAAGGGGAGTCTCTCGGATACAGC 2280
Db 2221 TTCGTGTGTGCGCGGTGAGCGGAGCTGTGTGTGAAGGGGAGTCTCTCGGATACAGC 2280
OY 2281 GGCAGGGGACAGCGGTGAGAGCTTGAAGGCTTCCCGGCCCATCTGTGAGCCGCTGAC 2340
Db 2281 GGCAGGGGACAGCGGTGAGAGCTTGAAGGCTTCCCGGCCCATCTGTGAGCCGCTGAC 2340
OY 2341 GTGAGGTCCTCTCCGTGGGGAAGTGAACCGCGCCGGGGTCCGTACTCTTCTATCTG 2400
Db 2341 GTGAGGTCCTCTCCGTGGGGAAGTGAACCGCGCCGGGGTCCGTACTCTTCTATCTG 2400
OY 2401 CCCAAAGAGCTCGGAGAGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTG 2460
Db 2401 CCCAAAGAGCTCGGAGAGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTG 2460
OY 2461 TTGCAACAACAGGTCTTCAAGCTTCTCAACCAAGTGTGAAGCGGAGCAGAGGCCCT 2520
Db 2461 TTGCAACAACAGGTCTTCAAGCTTCTCAACCAAGTGTGAAGCGGAGCAGAGGCCCT 2520
OY 2521 GCACGCTGGTGGTGGAGAGCTGGGGGGCGTGTCCGGGAGCTGGGGAGTGGGCTGTGAG 2580
Db 2521 GCACGCTGGTGGTGGAGAGCTGGGGGGCGTGTCCGGGAGCTGGGGAGTGGGCTGTGAG 2580
OY 2581 AAGCGGGCGGTGAGTGTGCGGGCTCCGCGGGAGCGCACAGTGTCTGTGATGCA 2640
Db 2581 AAGCGGGCGGTGAGTGTGCGGGCTCCGCGGGAGCGCACAGTGTCTGTGATGCA 2640
OY 2641 GCGCATGGGCGGTGAGACACAAAGCTTGGGGGAGCCCTCCCTGAGTGTGATGCA 2700
Db 2641 GCGCATGGGCGGTGAGACACAAAGCTTGGGGGAGCCCTCCCTGAGTGTGATGCA 2700
OY 2701 GCTTGTACACCTGTCTCAAGAGCTGCGCGGGGATTTAGAGCGGCTCACTCAAGTGT 2760
Db 2701 GCTTGTACACCTGTCTCAAGAGCTGCGCGGGGATTTAGAGCGGCTCACTCAAGTGT 2760
OY 2761 GTGGGCAAGAGGCGGCTGTGCGCGGGAGCAAGTGTGCAACCGGACCGCCAG 2820
Db 2761 GTGGGCAAGAGGCGGCTGTGCGCGGGAGCAAGTGTGCAACCGGACCGCCAG 2820
OY 2821 GAGCTGACCTTGTGCTGCTGAGGCGGCTGA 2853
Db 2821 GAGCTGACCTTGTGCTGCTGAGGCGGCTGA 2853

RESULT 5
US-60-453-135-7278
Sequence 7278, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARILLI, Michele
APPLICANT: IAKUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CI001456
CURRENT APPLICATION NUMBER: US/60/453,135
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7278
LENGTH: 2853
TYPE: DNA
ORGANISM: Homo sapiens
US-60-453-135-7278

|||||
Db 2101 GCCATCCCGCCAGCGGCTCAAGCATGACATCCGCCAGGCGGTATACAAAGGGCTGATY 2160
OY 2161 GGGGATGACAACTACCTGGGCTCTGAAGAAAGCAAGCAAGCAAGTACTGCTCAAGGGGCTAT 2220
Db 2161 GGGGATGACAACTACCTGGGCTCTGAAGAAAGCAAGCAAGTACTGCTCAAGGGGCTAT 2220
OY 2221 TTGCTGATGTCGGGCTGAGAGGCACTGTGTGTAAGAGCACTGTGCTGGGCTACAG 2280
Db 2221 TTGCTGATGTCGGGCTGAGAGGCACTGTGTGTAAGAGCACTGTGCTGGGCTACAG 2280
OY 2281 GGCACGCGCAGACGGGTGAGAGCCTGACAGGCTTCCCGGCCATCCCTGAGACCTGAC 2340
Db 2281 GGCACGCGCAGACGGGTGAGAGCCTGACAGGCTTCCCGGCCATCCCTGAGACCTGAC 2340
OY 2341 GTGAGAGTCCCTCCGCGGGGAGATGACACCGCGCGGCTCCGCTACTACTTCTATCTG 2400
Db 2341 GTGAGAGTCCCTCCGCGGGGAGATGACACCGCGCGGCTCCGCTACTACTTCTATCTG 2400
OY 2401 CCCAAGAGCCTCGGAGAGCAAGTCTCTCATCCCAAGAGCCTCGGAGACCTCTGTC 2460
Db 2401 CCCAAGAGCCTCGGAGAGCAAGTCTCTCATCCCAAGAGCCTCGGAGACCTCTGTC 2460
OY 2461 TTGCAACAAGCCTCTCAAGCTCTCAACAGGTGAGACAGCGGAGCAGAGGCCCT 2520
Db 2461 TTGCAACAAGCCTCTCAAGCTCTCAACAGGTGAGACAGCGGAGCAGAGGCCCT 2520
OY 2521 GCACGCTGGGCTGCGAGCTGAGGCGGCTGCTCCGAGCTGCGGAGCTGCGGAGCTGAG 2580
Db 2521 GCACGCTGGGCTGCGAGCTGAGGCGGCTGCTCCGAGCTGCGGAGCTGCGGAGCTGAG 2580
OY 2581 AAGCGGCGGTGAGCTGTCGCGGCTGCGGCGGAGCAGCGAGTCCCTGCTGTATGCA 2640
Db 2581 AAGCGGCGGTGAGCTGTCGCGGCTGCGGCGGAGCAGCGAGTCCCTGCTGTATGCA 2640
OY 2641 GCCATCGGCGGTGAGACAAAGCTGCGGAGAGCTGCGGAGCTGCGGAGCTGAG 2700
Db 2641 GCCATCGGCGGTGAGACAAAGCTGCGGAGAGCTGCGGAGCTGCGGAGCTGAG 2700
OY 2701 GCGTGTCAACCTGCTCAAGAGCTGCGGCGGAGTTCAGAGGCGCTCACTCAAGT 2760
Db 2701 GCGTGTCAACCTGCTCAAGAGCTGCGGCGGAGTTCAGAGGCGCTCACTCAAGT 2760
OY 2761 GTGGGCGCAGGAGC 2775
Db 2761 GTGGGCGCAGGAGC 2775

RESULT 6
US-60-453-050-7278
; Sequence 7278, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001457
; CURRENT APPLICATION NUMBER: US/60/453, 050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7278
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-7278

Query Match 86.5%; Score 2469; DB 11; Length 2853;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGCTTGTGTCGGGATCTTAACCTGGCTTTCGCGGCGCAACCGCTGAGGCTTGTAG 60
|||||

Db 1 ATGCTTGTGTCGGGATCTTAACCTGGCTTTCGCGGCGCAACCGCTGAGGCTTGTAG 60
OY 61 CCAGAGCGGAGAGTATGCTTCCATCCGAGCTGAGACCGCGGACATTACGGCGCGCTTAC 120
Db 61 CCAGAGCGGAGAGTATGCTTCCATCCGAGCTGAGACCGCGGACATTACGGCGCGCTTAC 120
OY 121 TACTGGCGGCGGCTCGGAGAGCTCCGCGGAGATCAGGAGCTATTTTCAGATACAGCAATT 180
Db 121 TACTGGCGGCGGCTCGGAGAGCTCCGCGGAGATCAGGAGCTATTTTCAGATACAGCAATT 180
OY 181 CAGGAGAGCTTTTACCTACCTGACCGCGGATGCTGATCTTGGCTCCGCGCTTCTG 240
Db 181 CAGGAGAGCTTTTACCTACCTGACCGCGGATGCTGATCTTGGCTCCGCGCTTCTG 240
OY 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTCAACGGGGCTCTTTCAGACTGCGAGC 300
Db 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTCAACGGGGCTCTTTCAGACTGCGAGC 300
OY 301 TGCCTCTATTCGCGGAGCTGTAAGCGCGGAGCTGCTGCGCTGTGAGGCTGTGC 360
Db 301 TGCCTCTATTCGCGGAGCTGTAAGCGCGGAGCTGCTGCGCTGTGAGGCTGTGC 360
OY 361 GGGGGCTCCGCGGAGCCTTGGCTACCGAGGCGCGGATGTCATTTAGCCGCTGCC 420
Db 361 GGGGGCTCCGCGGAGCCTTGGCTACCGAGGCGCGGATGTCATTTAGCCGCTGCC 420
OY 421 AATGCTACCGGCGGCGGCGGAGCGCAACAGCCAGGGCGCACTTCTTCAGCGCGG 480
Db 421 AATGCTACCGGCGGCGGCGGAGCGCAACAGCCAGGGCGCACTTCTTCAGCGCGG 480
OY 481 GGTGTTCCGGGCGGCGCTTCCGGAAGCCCACTCTGCTGCGGAGGCTGCGGCTG 540
Db 481 GGTGTTCCGGGCGGCGCTTCCGGAAGCCCACTCTGCTGCGGAGGCTGCGGCTG 540
OY 541 AACCCGCGCATCTACGGGCGCTTGACCCCTTACAAGCGCGGCGGCGGCTTCCAGGAG 600
Db 541 AACCCGCGCATCTACGGGCGCTTGACCCCTTACAAGCGCGGCGGCGGCTTCCAGGAG 600
OY 601 AGTGTACCGGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Db 601 AGTGTACCGGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
OY 661 GAGAGCGTGTGTGTCGCGGAGGAGTCAATGCTCAAGTTCACAGGCGGCGGCGGCGGAG 720
Db 661 GAGAGCGTGTGTGTCGCGGAGGAGTCAATGCTCAAGTTCACAGGCGGCGGCGGAG 720
OY 721 TATCTGTGAGCGTGTGTGCAAGCGGCGGCGGAGCTTACCGGCAATCCAGCATCTTCAAC 780
Db 721 TATCTGTGAGCGTGTGTGCAAGCGGCGGCGGAGCTTACCGGCAATCCAGCATCTTCAAC 780
OY 781 CCCATCAACATCGTTGTGTCAGGCTGCTGCTTGTAGATGCTGATCTGCGGCGGCGGAG 840
Db 781 CCCATCAACATCGTTGTGTCAGGCTGCTGCTTGTAGATGCTGATCTGCGGCGGCGGAG 840
OY 841 GTCAACCGGCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 841 GTCAACCGGCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
OY 901 AAATGTGATGACAGAGCAGCGGAGTACTGAGACAGTCCATCTCTTACAGGAGAG 960
Db 901 AAATGTGATGACAGAGCAGCGGAGTACTGAGACAGTCCATCTCTTACAGGAGAG 960
OY 961 CTGTGTGAGAGCACACCTGTGACACCTGCGGAGTGTGATGAGGCTTACCATGTGTGAC 1020
Db 961 CTGTGTGAGAGCACACCTGTGACACCTGCGGAGTGTGATGAGGCTTACCATGTGTGAC 1020
OY 1021 CCCAAGAGAGCTGCTCTCATTTAGAGAGGATGAGGCTTCCATCAGACTTCAACACTGCC 1080
Db 1021 CCCAAGAGAGCTGCTCTCATTTAGAGAGGATGAGGCTTCCATCAGACTTCAACACTGCC 1080
OY 1081 CAGGAGCTGGGCGCAGCTTCAACATGCCCCATGACATATGTGAAGTCTGTGAGAGGCTG 1140
Db 1081 CAGGAGCTGGGCGCAGCTTCAACATGCCCCATGACATATGTGAAGTCTGTGAGAGGCTG 1140

1141 TTGGAAGCTCGAGCCAAACACATGATGTCCTCCGACCTCATCCAGATCGACCGTGC 1200
1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCTCATCCAGATCGACCGTGC 1200
1201 AACCCCTGATCAGCTCAGTGTGCTGCATCATCAGCATCTTCTGAGACGCGGACGGT 1260
1201 AACCCCTGATCAGCTCAGTGTGCTGCATCATCAGCATCTTCTGAGACGCGGACGGT 1260
1261 GACGCTCCTGAGCAACCCAGACATCCCTGCGCCAGATGTCGCGGCGGC 1320
1261 GACGCTCCTGAGCAACCCAGACATCCCTGCGCCAGATGTCGCGGCGGC 1320
1321 AGCTACACCTGAGCCAGAGTGCAGTGTGCTGAGGCTTGGGCTGAGGCTGCTGCT 1380
1321 AGCTACACCTGAGCCAGAGTGCAGTGTGCTGAGGCTTGGGCTGAGGCTGCTGCT 1380
1381 TACATGACATGATCAGCAGAGTGTGCTGACCGGAGAGCCCAAGGACAGATGTGTGC 1440
1381 TACATGACATGATCAGCAGAGTGTGCTGACCGGAGAGCCCAAGGACAGATGTGTGC 1440
1441 CAGACCCGACATCTCCCTGGGCGGATGAGCAGCTGTGCGAGGCGAGCTGTGCTG 1500
1441 CAGACCCGACATCTCCCTGGGCGGATGAGCAGCTGTGCGAGGCGAGCTGTGCTG 1500
1501 AAAGGGGCTGCGTGGAGAGACACACCTCAACAGCAGGCGTGTGCTGCGGC 1560
1501 AAAGGGGCTGCGTGGAGAGACACACCTCAACAGCAGGCGTGTGCTGCGGC 1560
1561 AAATGGGATCCCTATGCGCCCTGTGCGCAGCATGTGTGCGGCGTGTGCGCAGG 1620
1561 AAATGGGATCCCTATGCGCCCTGTGCGCAGCATGTGTGCGGCGTGTGCGCAGG 1620
1621 AGGCAATGACCAACCCACCTGCGCAAGGAGGCGAAGTACTGCGAGGAGTGAAGGTG 1680
1621 AGGCAATGACCAACCCACCTGCGCAAGGAGGCGAAGTACTGCGAGGAGTGAAGGTG 1680
1681 AAATACCGATCTGCAATCTGAGAGCCTGCGCCAGCTCAGCTTCGSAAGAGCTTCGG 1740
1681 AAATACCGATCTGCAATCTGAGAGCCTGCGCCAGCTCAGCTTCGSAAGAGCTTCGG 1740
1741 GAGAGCAGTGTGAGGCTTTCACAGGCTACACACAGCAGCAGGCGTCACTTCGCGC 1800
1741 GAGAGCAGTGTGAGGCTTTCACAGGCTACACACAGCAGCAGGCGTCACTTCGCGC 1800
1801 GTGGCATGGGTGCCAAGTACTCGGGGTGTCTCCCGGAGCAAGTCAAGCTCATGTGC 1860
1801 GTGGCATGGGTGCCAAGTACTCGGGGTGTCTCCCGGAGCAAGTCAAGCTCATGTGC 1860
1861 CGAGCCATGAGCTGCTACTTCTATGTGCTGACACCAAGGAGTGTGAGCGGACGCTG 1920
1861 CGAGCCATGAGCTGCTACTTCTATGTGCTGACACCAAGGAGTGTGAGCGGACGCTG 1920
1921 TGTCTCTGATGCTGACCTGCTGCTGTGTGTCACAGGCAAGGCAAGGCTGCTGAT 1980
1921 TGTCTCTGATGCTGACCTGCTGCTGTGTGTCACAGGCAAGGCAAGGCTGCTGAT 1980
1921 TGTCTCTGATGCTGACCTGCTGCTGTGTGTCACAGGCAAGGCAAGGCTGCTGAT 1980
1981 GGAAGCCTGGGCTCCAGAAAGATGTGAGAAAGTGTGGGTGTGTGGGGAGACATTAAG 2040
1981 GGAAGCCTGGGCTCCAGAAAGATGTGAGAAAGTGTGGGTGTGTGGGGAGACATTAAG 2040
2041 AGCTGCAAGAGTGTGACTGTGCTTTCACCAAGCCATGCAATGCTCAATTTCTGTGTG 2100
2041 AGCTGCAAGAGTGTGACTGTGCTTTCACCAAGCCATGCAATGCTCAATTTCTGTGTG 2100
2101 GGCATCCCGGAGGCGCTCAAGCATGCAATGCGGAGGCGGTTCACAAAGGCGTATC 2160
2101 GGCATCCCGGAGGCGCTCAAGCATGCAATGCGGAGGCGGTTCACAAAGGCGTATC 2160
2161 GGGGATGACACTACTGTGCTGTGAAGAACAGCCAAAGGCAAGTACTGCTCAACGGGCT 2220
2161 GGGGATGACACTACTGTGCTGTGAAGAACAGCCAAAGGCAAGTACTGCTCAACGGGCT 2220

2221 TTCTGTGTGTGCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTGTGCTGCTACAGC 2280
2221 TTCTGTGTGTGCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTGTGCTGCTACAGC 2280
2281 GGCAGGCGACAGCGGTGGAGAGCTGCAAGCTTTCGCGGCCCATTCCTGAGACCCCTAAC 2340
2281 GGCAGGCGACAGCGGTGGAGAGCTGCAAGCTTTCGCGGCCCATTCCTGAGACCCCTAAC 2340
2341 GTGAGAGTCTCTGCGGTGGAGAGTACACCGCCCGGCTGCTGCTGCTGCTGCTGCTG 2400
2341 GTGAGAGTCTCTGCGGTGGAGAGTACACCGCCCGGCTGCTGCTGCTGCTGCTGCTG 2400
2401 CCCAAAGAGCTGCGGAGAGCAAGTCTCTCATATCCCAAGAACCCCGGAGACCTGTGTC 2460
2401 CCCAAAGAGCTGCGGAGAGCAAGTCTCTCATATCCCAAGAACCCCGGAGACCTGTGTC 2460
2461 TTGCACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
2461 TTGCACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
2521 GCAGCTGAGTGTGCGAGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
2521 GCAGCTGAGTGTGCGAGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
2581 AAGCGGCGGTGAGTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
2581 AAGCGGCGGTGAGTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
2641 GCGCATGCGCGGTGAGAGACACACCTGCGGAGGAGCCCTGACCTGGAGCTGAGC 2700
2641 GCGCATGCGCGGTGAGAGACACACCTGCGGAGGAGCCCTGACCTGGAGCTGAGC 2700
2701 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
2701 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
2761 GTGGGCGACGAGGCG 2775
2761 GTGGGCGACGAGGCG 2775

RESULT 7
US-10-311-035-32
; Sequence 32, Application US/10311035
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELIOT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAU, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANT, Angelo M.
; APPLICANT: BAUGHN, Mariah B.
; APPLICANT: NGUYEN, Danielle B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAMLA, Nandinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases

```
FILE REFERENCE: PI-0123 PCT
CURRENT APPLICATION NUMBER: US/10/311,035
CURRENT FILING DATE: 2002-12-10
PRIORITY APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
PRIORITY FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 32
LENGTH: 2930
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7473089CB1
US-10-311-035-32

Query Match      77.0%; Score 2196; DB 9; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGCTTGTGCTGGGATCTTACCTGATGCTGCTTGGCGGGGGAACCGCTGAGGCTGTAG 60
DB 75 ATGCTTGTGCTGGGATCTTACCTGATGCTGCTTGGCGGGGGAACCGCTGAGGCTGTAG 134
QY 61 CCAGAGCGGGAGGTAGTCTGCTTCCATCGACTGGACCCCGACATTAACCGCCGCTAC 120
DB 135 CCAGAGCGGGAGGTAGTCTGCTTCCATCGACTGGACCCCGACATTAACCGCCGCTAC 194
QY 121 TACTGGCGGGGTCGCGAGAGCTCCGGGATCAAGGAGCTATTTTCAATCAGAGATT 180
DB 195 TACTGGCGGGGTCGCGAGAGCTCCGGGATCAAGGAGCTATTTTCAATCAGAGATT 254
QY 181 CAGAGAGCTTTTACCTACCTGAGCGGAGATGCTCACTTGGCTCCCGCTTCTCC 240
DB 255 CAGAGAGCTTTTACCTACCTGAGCGGAGATGCTCACTTGGCTCCCGCTTCTCC 314
QY 241 ACTGAGACTGTGGGCTCCCTCCCTCAAGGGGCTACCGGGGGCTCTTCAGACCTGAGCG 300
DB 315 ACTGAGACTGTGGGCTCCCTCCCTCAAGGGGCTACCGGGGGCTCTTCAGACCTGAGCG 374
QY 301 TGGCTTATTTGGGGAGCTGAACGGCCGAGCCGAGACTGCTTGGCTGTGAGCTGTGC 360
DB 375 TGGCTTATTTGGGGAGCTGAACGGCCGAGCCGAGACTGCTTGGCTGTGAGCTGTGC 434
QY 361 GGGGGGCTCCGGGAGACCTTTGGCTACCGAGGCGCGAGATGATGATTAGCCGCTGCC 420
DB 435 GGGGGGCTCCGGGAGACCTTTGGCTACCGAGGCGCGAGATGATGATTAGCCGCTGCC 494
QY 421 AATGCTAGCGCGCGCGGCGGAGCGCAAGCGAGGGCGACACCTTCTCCAGCGCGG 480
DB 495 AATGCTAGCGCGCGCGGCGGAGCGCAAGCGAGGGCGACACCTTCTCCAGCGCGG 554
QY 481 GGTGTTCCGGGGGGGCTTCGGGAGACCCACCTCTCGCTGGGGGGTGGCTCGGGGCTGG 540
DB 555 GGTGTTCCGGGGGGGCTTCGGGAGACCCACCTCTCGCTGGGGGGTGGCTCGGGGCTGG 614
QY 541 AACCCCGCATCTTACCGGCGCTTGAACCTTACAAAGCCGCGGCGGGGCTTGGGGAG 600
DB 615 AACCCCGCATCTTACCGGCGCTTGAACCTTACAAAGCCGCGGCGGGGCTTGGGGAG 674
QY 601 AGTCGTAGCGCGCGGAGGTCTGGGCGCGCAAGCGCTTTCGTCTATCCCGGTACGTG 660
DB 675 AGTCGTAGCGCGCGGAGGTCTGGGCGCGCAAGCGCTTTCGTCTATCCCGGTACGTG 734
QY 661 GAGAGCGTGTGTGTGCGGAGAGTCAATGTCAGTTCAGCGGGGCGAGACTGGAAAT 720
DB 735 GAGAGCGTGTGTGTGCGGAGAGTCAATGTCAGTTCAGCGGGGCGAGACTGGAAAT 794
QY 721 TATGCTGACGCTGCTGGCAACGGCGGAGCTCTACCGGCATCCAGCATCTCTAAC 780
DB 795 TATGCTGACGCTGCTGGCAACGGCGGAGCTCTACCGGCATCCAGCATCTCTAAC 854
QY 781 CCCATCAACATCGTTGTGTGTAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
```

```
DB 855 CCCATCAACATCGTTGTGTGTAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 914
QY 841 GTACCCGCAATGGGCGCTTGAACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTGAA 900
DB 915 GTACCCGCAATGGGCGCTTGAACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTGAA 974
QY 901 AAATGAGTGAACACACCCGAGTACGAGGAGACCTGCACTCTTCCAGGACAGGAC 960
DB 975 AAATGAGTGAACACACCCGAGTACGAGGAGACCTGCACTCTTCCAGGAGAGGAC 1034
QY 961 CTGTGTGAGGACCCACCTGTGACACCTGGGAGCTGAGTGTGAGTGTGAGTGTGAG 1020
DB 1035 CTGTGTGAGGACCCACCTGTGACACCTGGGAGCTGAGTGTGAGTGTGAGTGTGAG 1094
QY 1021 CCCAAGAGAGCTGCTGTGATTTAGAGAGCATGGGCTTCCATCAGCTTCCACCTGCC 1080
DB 1095 CCCAAGAGAGCTGCTGTGATTTAGAGAGCATGGGCTTCCATCAGCTTCCACCTGCC 1154
QY 1081 CACGAGCTGGGGCAGGTGTCAACATGCCCATGACATGTGAAGTCTGTGAGAGAGTG 1140
DB 1155 CACGAGCTGGGGCAGGTGTCAACATGCCCATGACATGTGAAGTCTGTGAGAGAGTG 1214
QY 1141 TTTGGGAGCTCCGAGCCACACATGATGTCCCGAACCCCTCATCCAGATGAGCCG 1200
DB 1215 TTTGGGAGCTCCGAGCCACACATGATGTCCCGAACCCCTCATCCAGATGAGCCG 1274
QY 1201 AACCCCTGTGAGCTGTGATGCTGCTCCATCAATCAACCGACTTCTGGAACGCGGAGCG 1260
DB 1275 AACCCCTGTGAGCTGTGATGCTGCTCCATCAATCAACCGACTTCTGGAACGCGGAGCG 1334
QY 1261 GACTGCTCTGAGGAGCAACCGAGAGCCATCTCCCTGCCCGAGAGATCTGCGGGGCGC 1320
DB 1335 GACTGCTCTGAGGAGCAACCGAGAGCCATCTCCCTGCCCGAGAGATCTGCGGGGCGC 1394
QY 1321 AGCTACACCTGTGAGGAGCAACCGAGAGCTTGGCGTGGGCGCTCCAGACCTTGTCT 1380
DB 1395 AGCTACACCTGTGAGGAGCAACCGAGAGCTTGGCGTGGGCGCTCCAGACCTTGTCT 1454
QY 1381 TACATGACGTACTGACCAACAGCTGTGTGTCACCGGGAAGGCCAAGGAGAGATGTGTC 1440
DB 1455 TACATGACGTACTGACCAACAGCTGTGTGTCACCGGGAAGGCCAAGGAGAGATGTGTC 1514
QY 1441 CAGACCGGCACTTCCCTGGGGGAGTGGGAGCGAGCTGTGGCGAGGCAACCTTGTCTC 1500
DB 1515 CAGACCGGCACTTCCCTGGGGGAGTGGGAGCGAGCTGTGGCGAGGCAACCTTGTCTC 1574
QY 1501 AAAGGGGCTGCTGTGAGAGAGACACACCTCAACAGACAGGGGTGATGTTCTTGGGGC 1560
DB 1575 AAAGGGGCTGCTGTGAGAGAGACACACCTCAACAGAGGGGTGATGTTCTTGGGGC 1634
QY 1561 AAATGGGATCTTATGCGCCCTGCTGCGGACATGTGTGGGGGCTGACAGCTGGCCAG 1620
DB 1635 AAATGGGATCTTATGCGCCCTGCTGCGGACATGTGTGGGGGCTGACAGCTGGCCAG 1694
QY 1621 AGGAGTGAACCAACCCACCCCTGCGCAAGGGGGGCAAGTACGAGGGAGTGAAGGTG 1680
DB 1695 AGGAGTGAACCAACCCACCCCTGCGCAAGGGGGGCAAGTACGAGGGAGTGAAGGTG 1754
QY 1681 AAATACGATCTGCAATCTGAGACCTTGGCCAGCTGACCTCCGGAAGAGCTTCCG 1740
DB 1755 AAATACGATCTGCAATCTGAGACCTTGGCCAGCTGACCTCCGGAAGAGCTTCCG 1814
QY 1741 GAGGAGAGTGTAGGCTTTTCAACGGGTAAACACAGCAACCGGCTCACTTCCG 1800
DB 1815 GAGGAGAGTGTAGGCTTTTCAACGGGTAAACACAGCAACCGGCTCACTTCCG 1874
QY 1801 GTGCGATGGTGGCAAGTACTCGGGGTGTCTCCCGGGGAGCAAGTGAAGCTCATCTGC 1860
DB 1875 GTGCGATGGTGGCAAGTACTCGGGGTGTCTCCCGGGGAGCAAGTGAAGCTCATCTGC 1934
QY 1861 CGAGCAATGAGCACTGCTACTTCTATGTGCTGGACCCAA--GGTGTGAGAGGCGAG 1917
```



```
QY 1658 AGTACTGCGAGGAGTGAAGGTGAATACCATCTGCAATCTGGAGACCCCTGCCAGCT 1717
      |||||||
Db 61 AGTACTGCGAGGAGTGAAGGTGAATACCATCTGCAATCTGGAGACCCCTGCCAGCT 120
QY 1718 CACCTTCGCGAAAGAGCTTCGCGAGAGACAGTGTAGCTTTCAACGGCTCAACACACA 1777
      |||||||
Db 121 CACCTTCGCGAAAGAGCTTCGCGAGAGACAGTGTAGCTTTCAACGGCTCAACACACA 180
QY 1778 GCACCAACCGGCTCACTCTCG 1798
      |||||||
Db 181 GCACCAACCGGCTCACTCTCG 201

RESULT 11
US-60-453-135-77893
; Sequence 77893, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77893
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-77893

Query Match          5.3%; Score 150; DB 11; Length 201;
Best Local Similarity 99.5%; Pred. No. 5.1e-65;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1768 TACACACAGACCAACACCGGCTCACTCTGCGGTGCGATGSGGTGCCCAAGTACTCCGCG 1827
      |||||||
Db 1 TACACACAGACCAACACCGGCTCACTCTGCGGTGCGATGSGGTGCCCAAGTACTCCGCG 60
QY 1828 GTGTCTCCCGGGAGACAGTGAAGTCATCTGCGAGCCCAATGAGACAGTGTACTTTCTAT 1887
      |||||||
Db 61 GTGTCTCCCGGGAGACAGTGAAGTCATCTGCGAGCCCAATGAGACAGTGTACTTTCTAT 120
QY 1888 GTGTGCGACCCAGAGTGTGTGACGCGACGTGTCTCTCTGAGTCCACCTCCGCTGT 1947
      |||||||
Db 121 GTGTGCGACCCAGAGTGTGTGACGCGACGTGTCTCTCTGAGTCCACCTCCGCTGT 180
QY 1948 GTCCAAAGGAAGTGCATCAAG 1968
      |||||||
Db 181 GTCCAAAGGAAGTGCATCAAG 201

RESULT 12
US-60-453-135-77894
; Sequence 77894, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77894
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-77894

Query Match          5.3%; Score 150; DB 11; Length 201;
```

```
Best Local Similarity 99.5%; Pred. No. 5.1e-65;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2186 AGAACGCCAAGGCAAGTACTGCTCAACGGGCAATTCGTGTGTGCGCGGTGAGCGG 2245
      |||||||
Db 1 AGAACGCCAAGGCAAGTACTGCTCAACGGGCAATTCGTGTGTGCGCGGTGAGCGG 60
QY 2246 ACCTGTGTGAAGGCGAGTGTCTGCGGTACACGGGACGCGGACACAGCGGTGAGAGCC 2305
      |||||||
Db 61 ACCTGTGTGAAGGCGAGTGTCTGCGGTACACGGGACGCGGACACAGCGGTGAGAGCC 120
QY 2306 TGCAGGCTTCCCGGCCATCCTGAGCGCGTGCAGCGTGGAGGTCTCCGTGGGAGAGA 2365
      |||||||
Db 121 TGCAGGCTTCCCGGCCATCCTGAGCGCGTGCAGCGTGGAGGTCTCCGTGGGAGAGA 180
QY 2366 TGACACCGCCCGGGTCCGCT 2386
      |||||||
Db 181 TGACACCGCCCGGGTCCGCT 201

RESULT 13
US-60-453-050-77889
; Sequence 77889, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77889
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-77889

Query Match          5.3%; Score 150; DB 11; Length 201;
Best Local Similarity 99.5%; Pred. No. 5.1e-65;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 138 GGACTCCGGGGATGAGGAGCATTTTTCAGATCAGCATTTTCAGAGGACTTTTACCT 197
      |||||||
Db 1 GGACTCCGGGGATGAGGAGCATTTTTCAGATCAGCATTTTCAGAGGACTTTTACCT 60
QY 198 ACACCTGACGCGGATGCTCAGTTCTTGAGTCCCGCTTCTCCACTGAGCATTTGAGCT 257
      |||||||
Db 61 ACACCTGACGCGGATGCTCAGTTCTTGAGTCCCGCTTCTCCACTGAGCATTTGAGCT 120
QY 258 CCCCCCTCAGGGGCTCACCGGGGGCTTTTCAGACCTGCGACGCTCTTCTATTCTCAGGA 317
      |||||||
Db 121 CCCCCCTCAGGGGCTCACCGGGGGCTTTTCAGACCTGCGACGCTCTTCTATTCTCAGGA 180
QY 318 CGTGAAGCGGAGCGGAGCTC 338
      |||||||
Db 181 CGTGAAGCGGAGCGGAGCTC 201

RESULT 14
US-60-453-050-77890
; Sequence 77890, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 19:52:06 ; Search time 5576 Seconds

(without alignments)
12664.332 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853 1 atgcctctgctggagcatcct.....gcgtccgagggctgctga 2853

Sequence: 1 atgcctctgctggagcatcct.....gcgtccgagggctgctga 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents, NA, Main:*

```

1: /cgn2_6/ptodata/1/pna/US0955A_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US0955B_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US0955C_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US0955D_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US0955E_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US0955F_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US0955G_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US0955H_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US0955I_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US0955J_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US0955K_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US0955L_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US0955M_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US0955N_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US0955O_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US0955P_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US0955Q_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US0955R_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US0955S_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US0955T_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US0955U_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US0955V_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US0955W_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US0955X_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US0955Y_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US0955Z_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US0956A_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US0956B_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US0956C_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US0956D_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US0956E_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US0956F_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US0956G_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US0956H_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US0956I_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US0956J_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US0956K_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US0956L_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US0956M_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US0956N_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US0956O_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US0956P_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US0956Q_COMB.seq:*

```

Result No.	Score	Query	Length	DB ID	Description
1	2853	100.0	2853	US-09-965-631-3	Sequence 3, Appl1
2	2853	100.0	3446	US-09-965-631-7	Sequence 7, Appl1
3	2848.2	99.8	2867	US-09-741-151-1	Sequence 1, Appl1
4	2846.6	99.8	2853	US-10-009-332-2	Sequence 2, Appl1
5	2817.6	98.8	2930	US-60-216-821-9*	Sequence 9, Appl1
6	2619.2	91.8	2804	US-10-093-463-27	Sequence 27, Appl1
7	2299.4	80.6	5714	US-60-360-207-12354	Sequence 12354, A
8	2298	80.5	2469	US-10-163-316-3	Sequence 3, Appl1
9	2298	80.5	2469	US-60-297-863-3	Sequence 3, Appl1
10	2298	80.5	2940	US-10-163-316-1	Sequence 1, Appl1
11	2298	80.5	2940	US-60-297-863-1	Sequence 1, Appl1
12	2285.4	80.1	3400	US-60-212-656-762	Sequence 762, App
13	2285.4	78.0	3759	US-60-242-679-1814	Sequence 1814, Ap
14	1997.8	66.9	3635	US-60-230-435-2152	Sequence 2152, Ap
15	1671.4	58.6	2297	US-10-093-463-29	Sequence 29, Appl1
16	1285.8	45.1	1829	US-09-575-003-17	Sequence 17, Appl1
17	1285.8	45.1	1829	US-09-609-059-2	Sequence 2, Appl1
18	1285.8	45.1	1829	US-09-975-545-17	Sequence 17, Appl1
19	1285.8	45.1	1829	US-10-071-241-2	Sequence 2, Appl1
20	1091.2	38.2	1104	US-09-965-631-5	Sequence 5, Appl1
21	960.4	33.7	27392	US-60-206-028-17	Sequence 17, Appl1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```
22 960.4 33.7 27392 64 US-60-206-028-18 Sequence 18, Appl
23 960.4 33.7 27760 68 US-60-242-679-304 Sequence 304, App
24 960.4 33.7 28854 29 US-09-741-151-3 Sequence 3, Appl
25 960.4 33.7 31624 64 US-60-207-315-28 Sequence 28, Appl
26 960.4 33.7 31627 65 US-60-212-656-232 Sequence 232, Appl
27 960.4 33.7 32768 64 US-60-207-315-29 Sequence 29, Appl
28 960.4 33.7 32768 64 US-60-208-020-14 Sequence 14, Appl
29 960.4 33.7 32768 64 US-60-209-043-17 Sequence 17, Appl
30 960.4 33.7 32768 64 US-60-209-043-18 Sequence 18, Appl
31 960.4 33.7 32768 64 US-60-209-043-15 Sequence 19, Appl
32 960.4 33.7 32768 65 US-60-212-656-233 Sequence 233, Appl
33 960.4 33.7 32768 67 US-60-230-435-158 Sequence 158, App
34 960.4 33.7 32768 67 US-60-230-435-159 Sequence 159, App
35 959 33.6 966 36 US-09-965-631-1 Sequence 1, Appl
36 949.4 33.3 31959 64 US-60-200-382-27 Sequence 27, Appl
37 949.4 33.3 31959 64 US-60-200-382-28 Sequence 28, Appl
38 777.2 27.2 1875 64 US-60-207-213-37 Sequence 37, Appl
39 777.2 27.2 1875 65 US-60-213-845-8 Sequence 8, Appl
40 744.2 26.1 25614 64 US-60-206-028-19 Sequence 19, Appl
41 744.2 26.1 25614 64 US-60-206-028-20 Sequence 20, Appl
42 688.4 24.1 741 65 US-60-213-845-538 Sequence 538, App
43 664.4 23.3 3008 1 PCT-US00-14462A-3 Sequence 3, Appl
44 664.4 23.3 3008 1 PCT-US99-01313-3 Sequence 3, Appl
45 664.4 23.3 3008 16 US-09-235-810-3 Sequence 3, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-965-631-3
; Sequence 3, Application US/09965631
; GENERAL INFORMATION:
; APPLICANT: Frittdle, Carl Johan
; APPLICANT: Hildun, Elin
; FILE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; TITLE REFERENCE: Lex-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-3
```

```
Query Match 100.0%; Score 2853; DB 36; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCTTCTGCTGGGCACTCAACCCGCTTTGGCCGGCGGACCGCTGAGGCTCTGAG 60
|
Db 1 ATGCTTCTGCTGGGCACTCAACCCGCTTTGGCCGGCGGACCGCTGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGAGTAGTCTTCCATCCGACTGGACCCGGGACATTAACGCCGCCGCTAC 120
|
Db 61 CCAGAGCGGAGAGTAGTCTTCCATCCGACTGGACCCGGGACATTAACGCCGCCGCTAC 120
QY 121 TACTGCGGGGTCGCCGAGACTCCGGGATCAGGAGCTATTTTTCAGATCAGCATTT 180
|
Db 121 TACTGCGGGGTCGCCGAGACTCCGGGATCAGGAGCTATTTTTCAGATCAGCATTT 180
QY 181 CAGAGAGCTTTTACCTACACCTGAGCGCGGATGCTAGTTTGGCTCCGCTTCTGC 240
|
Db 181 CAGAGAGCTTTTACCTACACCTGAGCGCGGATGCTAGTTTGGCTCCGCTTCTGC 240
QY 241 ACTGAGCATCTGGGCTGCCCTCTCCAGAGGCTCACCGGGGCTTTCAGACTGCGAGCG 300
|
Db 241 ACTGAGCATCTGGGCTGCCCTCTCCAGAGGCTCACCGGGGCTTTCAGACTGCGAGCG 300
```

```
QY 301 TGTCTTATTCTGGGAGCGTGAACGCCGAGCTGTTCCGCTGCTGTGAGCTGTGCG 360
|
Db 301 TGTCTTATTCTGGGAGCGTGAACGCCGAGCTGTTCCGCTGCTGTGAGCTGTGCG 360
QY 361 GGGGGGCTCCGCGAGGCTTTGGCTTACCGAGGCGCGAGTATGTCATTAACCCGCTGCC 420
|
Db 361 GGGGGGCTCCGCGAGGCTTTGGCTTACCGAGGCGCGAGTATGTCATTAACCCGCTGCC 420
QY 421 AATCTAGACCGCGCGCGCGCGCGAGCGGACAGGCGGCGACACTTTCACGCGCGG 480
|
Db 421 AATCTAGACCGCGCGCGCGCGCGAGCGGACAGGCGGCGACACTTTCACGCGCGG 480
QY 481 GGTGTTCGCGGCGCGGCTTCCGAGACCCCACTCTGCTGCGGGGTGGCTCGGCTGCG 540
|
Db 481 GGTGTTCGCGGCGCGGCTTCCGAGACCCCACTCTGCTGCGGGGTGGCTCGGCTGCG 540
QY 541 AACCOCGCATCTTACGCGGCTTGGACCTTACAGCGCGCGCGCGCGGCTTCCGAGAG 600
|
Db 541 AACCOCGCATCTTACGCGGCTTGGACCTTACAGCGCGCGCGCGGCTTCCGAGAG 600
QY 601 AGTGTACCGCGGCGAGGCTGGGCGCGGCAAGCGTTTGTCTATCCCGGATACGTTG 660
|
Db 601 AGTGTACCGCGGCGAGGCTGGGCGCGGCAAGCGTTTGTCTATCCCGGATACGTTG 660
QY 661 GAGAGCGTGTGTGCGCGAGCTCAATGTCATGTCATGTCATGTCATGTCATGTCATG 720
|
Db 661 GAGAGCGTGTGTGCGCGAGCTCAATGTCATGTCATGTCATGTCATGTCATGTCATG 720
QY 721 TATCTGCTGAGCGCTGTGCGAGCGGCGGCGGCTTACCGCCATCCGACATCTCTAAC 780
|
Db 721 TATCTGCTGAGCGCTGTGCGAGCGGCGGCGGCTTACCGCCATCCGACATCTCTAAC 780
QY 781 CCCATCAACATGTTGTGTCAGAGTGTGCTTTAGATGTCGATGTCGATGTCGATGTCG 840
|
Db 781 CCCATCAACATGTTGTGTCAGAGTGTGCTTTAGATGTCGATGTCGATGTCGATGTCG 840
QY 841 GTACCGGCAATGCGGCGCTTACGCTGCGGCACTTGTGCTGCGGCAAGAACTGTAAC 900
|
Db 841 GTACCGGCAATGCGGCGCTTACGCTGCGGCACTTGTGCTGCGGCAAGAACTGTAAC 900
QY 901 AAAGTGAAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 960
|
Db 901 AAAGTGAAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 960
QY 961 CTGTGTGAGCGCACCACTGTGACACCTGCGGATGCTATGTGTGTGTGTGTGTGTGT 1020
|
Db 961 CTGTGTGAGCGCACCACTGTGACACCTGCGGATGCTATGTGTGTGTGTGTGTGTGT 1020
QY 1021 CCCAGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
|
Db 1021 CCCAGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1081 CACGAGCTGGGCGACGCTGTCAATGATGCGCCATGATGAAAGTGTGTGTGTGTGTGT 1140
|
Db 1081 CACGAGCTGGGCGACGCTGTCAATGATGCGCCATGATGAAAGTGTGTGTGTGTGTGT 1140
QY 1141 TTTGGAAAGCTCCGAGCCAAACCAATGATGTCGCGGACCTTCCATGATGATGATGATG 1200
|
Db 1141 TTTGGAAAGCTCCGAGCCAAACCAATGATGTCGCGGACCTTCCATGATGATGATGATG 1200
QY 1201 AACCCCTGTGTGAGCTGTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 1260
|
Db 1201 AACCCCTGTGTGAGCTGTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCTGCGCGAGGATCTCTCGGCGCGC 1320
|
Db 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCTGCGCGAGGATCTCTCGGCGCGC 1320
QY 1321 AACTACACCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
|
Db 1321 AACTACACCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 TACATGACGATCTGACCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
|
Db 1381 TACATGACGATCTGACCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
```


Db 697 TGCCTCTATTCTGGGAGCTGAGCCGAGCCGAGCTGTTCCGCTGCTGAGCGCTGTC 756
QY 361 GGGGGGCTCCGGGAGACCTTTGGCTACGAGGCGCCGAGTATGTCATTATATCCGCTGCC 420
Db 757 GGGGGGCTCCGGGAGACCTTTGGCTACGAGGCGCCGAGTATGTCATTATATCCGCTGCC 816
QY 421 AATGCTAGCGCGCGCGCGCGCGAGCGCAACAGCGCGCGCGCGCGCTTCCACGCGCG 480
Db 817 AATGCTAGCGCGCGCGCGCGCGAGCGCAACAGCGCGCGCGCGCGCTTCCACGCGCG 876
QY 481 GGTGTTCCGGCGCGCGCTTCCGAGAACCCACCTCTCGCTCGGCGGTGCGCTGCGGCTG 540
Db 877 GGTGTTCCGGCGCGCGCTTCCGAGAACCCACCTCTCGCTCGGCGGTGCGCTGCGGCTG 936
QY 541 AACCCCGCATCTACGCGCGCGCTTCCGAGAACCCGCGCGCGCGCGCTTCCGAGAAC 600
Db 937 AACCCCGCATCTACGCGCGCGCTTCCGAGAACCCGCGCGCGCGCGCTTCCGAGAAC 996
QY 601 AGTCGTAGCGCGCGCGAGCTGCGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 660
Db 997 AGTCGTAGCGCGCGCGAGCTGCGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 1056
QY 661 GAGACGCTGCTGCTGCGCGAGCTCATATGTCAGTTCAGCGCGCGCGCGCTTCCGAGAAC 720
Db 1057 GAGACGCTGCTGCTGCGCGAGCTCATATGTCAGTTCAGCGCGCGCGCGCTTCCGAGAAC 1116
QY 721 TATGCTAGCGCTGCTGCGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 780
Db 1117 TATGCTAGCGCTGCTGCGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 1176
QY 781 CCCATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 1177 CCCATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1236
QY 841 GTCACCGCGCAATGCGCGCGCTGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 900
Db 1237 GTCACCGCGCAATGCGCGCGCTGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 1296
QY 901 AAGTGAAGTACAGACCGCGCGCGCTGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 960
Db 1297 AAGTGAAGTACAGACCGCGCGCGCTGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 1356
QY 961 CTGTGTGAGCGCACCACTGTGACACCGCTGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 1020
Db 1357 CTGTGTGAGCGCACCACTGTGACACCGCTGAGCGCGCGCGCGCGCTTCCGAGAACCCGCGCGCG 1416
QY 1021 CCCAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1417 CCCAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476
QY 1081 CAGAGCTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1477 CAGAGCTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1536
QY 1141 TTTTGGAGAGTCCGAGCGCAACACATGATGTCCCGACCTCTCATCCAGATGCGAGCGTGGC 1200
Db 1537 TTTTGGAGAGTCCGAGCGCAACACATGATGTCCCGACCTCTCATCCAGATGCGAGCGTGGC 1596
QY 1201 AACCCCGCATCTACGCGCGCGCTTCCGAGAACCCGCGCGCGCGCGCTTCCGAGAAC 1260
Db 1597 AACCCCGCATCTACGCGCGCGCTTCCGAGAACCCGCGCGCGCGCGCTTCCGAGAAC 1656
QY 1261 GACTGCTCTCTGAGCAACCCAGACGCGCTTCCGAGAACCCGCGCGCGCGCGCTTCCGAGAAC 1320
Db 1657 GACTGCTCTCTGAGCAACCCAGACGCGCTTCCGAGAACCCGCGCGCGCGCGCTTCCGAGAAC 1716
QY 1321 AGCTACACCTGAGCGCGAGCTGCGAGCTGCTTGGCGTGGCGTCCAGACCGCTGCTGCTG 1380
Db 1717 AGCTACACCTGAGCGCGAGCTGCGAGCTGCTTGGCGTGGCGTCCAGACCGCTGCTGCTG 1776
QY 1381 TACATGAGATGCTGACACAGCTGTGTCACCGGAGAGCGCAAGGAGACAGATGCTGCTG 1440
Db 1777 TACATGAGATGCTGACACAGCTGTGTCACCGGAGAGCGCAAGGAGACAGATGCTGCTG 1836

QY 1441 CAGACCGCGCACTTCCCTGGCGCGAGTGGACAGAGCTGTGGCGAGGCGCAAGCTTGGCTG 1500
Db 1837 CAGACCGCGCACTTCCCTGGCGCGAGTGGACAGAGCTGTGGCGAGGCGCAAGCTTGGCTG 1896
QY 1501 AAGGGGCGCTGCTGAGAGACCAACCTTCACACAGCAGAGGCTGATGCTTCTGCGGC 1560
Db 1897 AAGGGGCGCTGCTGAGAGACCAACCTTCACACAGCAGAGGCTGATGCTTCTGCGGC 1956
QY 1561 AATGAGATCCATGCGCGCGCTGCGCGACATGCTGCTGCGCGCGCTGAGCTGACAG 1620
Db 1957 AATGAGATCCATGCGCGCGCTGCGCGACATGCTGCTGCGCGCGCTGAGCTGAGCTG 2016
QY 1621 AGCAGTGCACCAACCCGCGCGCGCTGCGCAACGCGCGCGCGCGCTGCGAGGAGTGAAGCTG 1680
Db 2017 AGCAGTGCACCAACCCGCGCGCGCTGCGCAACGCGCGCGCGCGCTGCGAGGAGTGAAGCTG 2076
QY 1681 AATATCCATCTGCAATCTGAGAGCGCTTCCGAGAACCCGCGCGCGCGCTTCCGAGAAC 1740
Db 2077 AATATCCATCTGCAATCTGAGAGCGCTTCCGAGAACCCGCGCGCGCGCTTCCGAGAAC 2136
QY 1741 GAGGAGAGTGTGAGGCTTCCAGCGCGCTGACAGCAGCAACCGCGCTGCTGCGC 1800
Db 2137 GAGGAGAGTGTGAGGCTTCCAGCGCGCTGACAGCAGCAACCGCGCTGCTGCGC 2196
QY 1801 GTGCGATGGTGGCCAAAGTACTCCGCGCTGCTTCCCGGAGCAAGTGCATGCTGCTG 1860
Db 2197 GTGCGATGGTGGCCAAAGTACTCCGCGCTGCTTCCCGGAGCAAGTGCATGCTGCTG 2256
QY 1861 CGAGCGAATGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 2257 CGAGCGAATGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2216
QY 1921 TGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 2317 TGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376
QY 1981 GGGAACTGGCTTCCAGAGAGATGAGCAAGTGTGGGCTGCTGAGAGAGCAATTAAG 2040
Db 2377 GGGAACTGGCTTCCAGAGAGATGAGCAAGTGTGGGCTGCTGAGAGAGCAATTAAG 2436
QY 2041 AGCTGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2437 AGCTGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2496
QY 2101 GCGATCCCGCGAGCGCTTCAAGCATGACATCCGCGCGCGCTTCCAGAGAGCAATTAAG 2160
Db 2497 GCGATCCCGCGAGCGCTTCAAGCATGACATCCGCGCGCGCTTCCAGAGAGCAATTAAG 2556
QY 2161 GGGGATGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2557 GGGGATGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2616
QY 2221 TTTGCTGTGTGGCGCGGTGAGCGCGAGCTGTGTGAAGGCGCATGTGCTGCTGCTGCTG 2280
Db 2617 TTTGCTGTGTGGCGCGGTGAGCGCGAGCTGTGTGAAGGCGCATGTGCTGCTGCTGCTG 2676
QY 2281 GGCACGCGGACAGCGGTGAGAGCGCTGAGGCTTCCGCGCGCGCTTCCAGAGAGCAATTAAG 2340
Db 2677 GGCACGCGGACAGCGGTGAGAGCGCTGAGGCTTCCGCGCGCGCTTCCAGAGAGCAATTAAG 2736
QY 2341 GTGAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2737 GTGAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2796
QY 2401 CCCAAGAGCGCTGCGAGAGCAAGTCTCTCATCCCAAGAGAGCGCGCGCGCGCTTCCGAGAAC 2460
Db 2797 CCCAAGAGCGCTGCGAGAGCAAGTCTCTCATCCCAAGAGAGCGCGCGCGCGCTTCCGAGAAC 2856
QY 2461 TTGACACAGAGCTCTGAGCTTCCCAACAGAGTGTGAGAGCGCGAGAGAGAGAGAGAGAG 2520
Db 2857 TTGACACAGAGCTCTGAGCTTCCCAACAGAGTGTGAGAGCGCGAGAGAGAGAGAGAGAGAG 2916

```

OY 2521 GCACGCTGGTGGCTGGCAGCTGGAGGCGCTGCTCCGAGCTCGGCACCTGGCTTCAG 2580
    |||
Db 2917 GCACGCTGGTGGCTGGCAGCTGGAGGCGCTGCTCCGAGCTCGGCACCTGGCTTCAG 2976
OY 2581 AAGCGGGGGGAGTGTGCGGGGCTCGCGCGGAGCGCACGGTCCCTGCTGTGATGCA 2640
    |||
Db 2977 AAGCGGGGGGAGTGTGCGGGGCTCGCGCGGAGCGCACGGTCCCTGCTGTGATGCA 3036
OY 2641 GCCCATGCGCCCTGGAGACACAAAGCCTGCGGGAGCGCTGCCACCTGGAGAGTACG 2700
    |||
Db 3037 GCCCATGCGCCCTGGAGACACAAAGCCTGCGGGAGCGCTGCCACCTGGAGAGTACG 3096
OY 2701 GCTTGTCACCTGCTCCAGAGCTGGCGCGGGGATTTCAGAGCGCTGCTCAAGTGT 2760
    |||
Db 3097 GCTTGTCACCTGCTCCAGAGCTGGCGCGGGGATTTCAGAGCGCTGCTCAAGTGT 3156
OY 2761 GTGGGCGAGGAGCGGCTGCTGGGCGGGAGCAGTGCACCTGCGACCGAGCCCGAG 2820
    |||
Db 3157 GTGGGCGAGGAGCGGCTGCTGGGCGGGAGCAGTGCACCTGCGACCGAGCCCGAG 3216
OY 2821 GAGCTGACTTCTGCGTCTGAGGCGGCTGCTGA 2853
    |||
Db 3217 GAGCTGACTTCTGCGTCTGAGGCGGCTGCTGA 3249

```

RESULT 3

```

US-09-741-151-1
; Sequence 1, Application US/09741151
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shiaoiping et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1001005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Human
US-09-741-151-1

```

```

Query Match          99.8%; Score 2848.2; DB 29; Length 2867;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2850; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 ATGCTTGTGCTGGCATCTTAACCTGGCTTCCGCGGCGACCGCTGAGGCTTGAG 60
    |||
Db 4 ATGCTTGTGCTGGCATCTTAACCTGGCTTCCGCGGCGACCGCTGAGGCTTGAG 63
OY 61 CCAGAGCGGAGAGTAGTGTGCTCCATCGACTGAGACCGGACATTAAAGGCGCGCTAC 120
    |||
Db 64 CCAGAGCGGAGAGTAGTGTGCTCCATCGACTGAGACCGGACATTAAAGGCGCGCTAC 123
OY 121 TACTGCGGGGGTCCGAGAGACTCGGGGATCAGGACTCATTTTTCAGATCAAGCAATT 180
    |||
Db 124 TACTGCGGGGGTCCGAGAGACTCGGGGATCAGGAGCACTTTTTCAGATCAAGCAATT 183
OY 181 CAGGAGAGATTTCCTACACTGACGCGGATGCTGCTGCTGCTGCTGCTTTC 240
    |||
Db 184 CAGGAGAGATTTCCTACACTGACGCGGATGCTGCTGCTGCTGCTGCTTTC 243
OY 241 ACTGAGCATCTGGCGTCCCTCCAGGGGCTCACCGGGGCTTTACAGACTGCGAGCG 300
    |||
Db 244 ACTGAGCATCTGGCGTCCCTCCAGGGGCTCACCGGGGCTTTACAGACTGCGAGCG 303
OY 301 TGCCTTATTTGGGAGAGTGAAGCGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
    |||
Db 304 TGCCTTATTTGGGAGAGTGAAGCGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
OY 361 GGGGGGCTCCGCGAGCCTTTGGCTACCGAGGCGCGAGTAGTGTGCTGCTGCTGCTGCTGCTGCTG 420

```

```

Db 364 GGGGGGCTCCGCGAGAGCTTTGGCTACCGAGGCGCGAGATGTGATTAAGCCGCTGCC 423
    |||
OY 421 AATGTACAGCGCGCGCGGCGACAGCGACAGGCGCGACACCTTCACAGCGCGCG 480
    |||
Db 424 AATGTACAGCGCGCGCGGCGACAGCGACAGGCGCGACACCTTCACAGCGCGCG 483
OY 481 GGTGTTCGCGGGGCGCTTCCGAGAGCCCAAGCTCTGCTGCGGGGTGAGCTGAGCTG 540
    |||
Db 484 GGTGTTCGCGGGGCGCTTCCGAGAGCCCAAGCTCTGCTGCGGGGTGAGCTGAGCTG 543
OY 541 AACCCGCGCATCTACAGGCGCTTGAGCCCTTACAGCGCGCGGCGGCGCTTCCGAGAG 600
    |||
Db 544 AACCCGCGCATCTACAGGCGCGCTTGAGCCCTTACAGCGCGCGGCGGCGCTTCCGAGAG 603
OY 601 AGTGTACCGGGCGAGGTCTGGGCGGCGCAAGGCTTGTGCTATGCTGCTGCTGCTGCTGCTG 660
    |||
Db 604 AGTGTACCGGGCGAGGTCTGGGCGGCGCAAGGCTTGTGCTATGCTGCTGCTGCTGCTGCTG 663
OY 661 GAGAGCTGTGTGTGCGGAGAGTCAATGATCAAGTTCACAGCGCGGAGCTGGAACAT 720
    |||
Db 664 GAGAGCTGTGTGTGCGGAGAGTCAATGATCAAGTTCACAGCGCGGAGCTGGAACAT 723
OY 721 TATCTGTGAGCGTGTGGCAAGCGGCGGAGCTTACCGCCATCCAGCATCTTCAC 780
    |||
Db 724 TATCTGTGAGCGTGTGGCAAGCGGCGGAGCTTACCGCCATCCAGCATCTTCAC 783
OY 781 CCCATCAACATCTGTTGTGTCAGGTGCTGCTCTTGAAGATCGTGACTCGGCGCCAG 840
    |||
Db 784 CCCATCAACATCTGTTGTGTCAGGTGCTGCTCTTGAAGATCGTGACTCGGCGCCAG 843
OY 841 GTACAGCGGCAATGCGGCGCTGAGCTGCGCAACTTGTGCTGCGGAGAGAACTGAGAC 900
    |||
Db 844 GTACAGCGGCAATGCGGCGCTGAGCTGCGCAACTTGTGCTGCGGAGAGAACTGAGAC 903
OY 901 AAAGTGTGACAAAGACCCCGAGTACTGGAGACTGCCATCTCTTACACAGGTAAGAC 960
    |||
Db 904 AAAGTGTGACAAAGACCCCGAGTACTGGAGACTGCCATCTCTTACACAGGTAAGAC 963
OY 961 CTGTGTGAGGCGACCACTGTGACACCGTGGGCTGATGTGGGTATCATGTGAGT 1020
    |||
Db 964 CTGTGTGAGGCGACCACTGTGACACCGTGGGCTGATGTGGGTATCATGTGAGT 1023
OY 1021 CCCAAGAGAGGTGCTGTGTCATTGAGAGAGTGGCTTCCATCAGCTTACCACTGCC 1080
    |||
Db 1024 CCCAAGAGAGGTGCTGTGTCATTGAGAGAGTGGCTTCCATCAGCTTACCACTGCC 1083
OY 1081 CAGGAGCTGGGCGACGCTTCAACATGCCCATGAGAAAGTGTGAGAGGCTG 1140
    |||
Db 1084 CAGGAGCTGGGCGACGCTTCAACATGCCCATGAGAAAGTGTGAGAGGCTG 1143
OY 1141 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGAGCCCTCATCGATGAGACGCTGCC 1200
    |||
Db 1144 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGAGCCCTCATCGATGAGACGCTGCC 1203
OY 1201 AACCCGCTGACGCTGTGAGAGTGTCCATCATACAGCACTTCGAGACAGCGGCGAGCT 1260
    |||
Db 1204 AACCCGCTGACGCTGTGAGAGTGTCCATCATACAGCACTTCGAGACAGCGGCGAGCT 1263
OY 1261 GACTGCTCTGTGAGCAACCCAGCAAGGCCATCTCCCTGCGCGAGATGTGCGGGCGCT 1320
    |||
Db 1264 GACTGCTCTGTGAGCAACCCAGCAAGGCCATCTCCCTGCGCGAGATGTGCGGGCGCT 1323
OY 1321 AGCTACACCTGTAGGACAGCATGTGAGGTGCTTTGGCGTGGGCTCCAGCCCTGTGCT 1380
    |||
Db 1324 AGCTACACCTGTAGGACAGCATGTGAGGTGCTTTGGCGTGGGCTCCAGCCCTGTGCT 1383
OY 1381 TACATGAGTACTGTGACCAAGCTGTGTGACACCGGAGGCGCAAGGAGAGTGTGCTG 1440
    |||
Db 1384 TACATGAGTACTGTGACCAAGCTGTGTGACACCGGAGGCGCAAGGAGAGTGTGCTG 1443
OY 1441 CAGACCGCGCACTTCCCTGGGCGGAGTGGACACAGCTGTGGCGAGGCAAGCTGTGCTG 1500
    |||

```


Db 1444 CAGACCCGCTCTCCCTCCGCGCATGGCACACGCTGTGGCGAGGCACTCTGCTC 1503
 Oy 1501 AAGGGGCTGGGTGGAGAGACAACTCAACAAACAGAGGGTGGATGTTCTCTGGGCC 1560
 Db 1504 AAGGGGCTGGGTGGAGAGACAACTCAACAAACAGAGGGTGGATGTTCTCTGGGCC 1563
 Oy 1561 AATGGGATCCCTATGAGCCCTGCTGCGGACATGTGGTGGGGGCGTGCAGCTGGCCAGG 1620
 Db 1564 AATGGGATCCCTATGAGCCCTGCTGCGGACATGTGGTGGGGGCGTGCAGCTGGCCAGG 1623
 Oy 1621 AGCAGTGCACCAACCCACCCCTGCAACAGGGGGCAAGTACTGCGAGGAGAGAGGTTG 1680
 Db 1624 AGCAGTGCACCAACCCACCCCTGCAACAGGGGGCAAGTACTGCGAGGAGAGAGGTTG 1683
 Oy 1681 AATATCCGATCTCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCGGAAAGAGTTCCGG 1740
 Db 1684 AATATCCGATCTCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCGGAAAGAGTTCCGG 1743
 Oy 1741 GAGAGCAGTGTGAGGCTTTCAAGGGCTACAAACAGACAGACACCGGCTCACTCTCGCC 1800
 Db 1744 GAGAGCAGTGTGAGGCTTTCAAGGGCTACAAACAGACAGACACCGGCTCACTCTCGCC 1803
 Oy 1801 GTGGCATGGGTGCCAAGTACTCCGGCTGTCTCCCGGGAGCAAGTGAAGCTCATCTGC 1860
 Db 1804 GTGGCATGGGTGCCAAGTACTCCGGCTGTCTCCCGGGAGCAAGTGAAGCTCATCTGC 1863
 Oy 1861 CGAGCCATGGGCACTGGCTACTTCTATGTGTGTGGACCCAAAGTGGTGGACGGGACGCTG 1920
 Db 1864 CGAGCCATGGGCACTGGCTACTTCTATGTGTGTGGACCCAAAGTGGTGGACGGGACGCTG 1923
 Oy 1921 TGGCTTCCTACTTCACACTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
 Db 1924 TGGCTTCCTACTTCACACTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1983
 Oy 1981 GGAAGCTGGGCTCCCAAGAGAGATTTGACAGTGTGGGGTGTGTGGGGAGACATTAAG 2040
 Db 1984 GGAAGCTGGGCTCCCAAGAGAGATTTGACAGTGTGGGGTGTGTGGGGAGACATTAAG 2043
 Oy 2041 AGCTGGCAAGAGGAGTGTGAGTCTTCAACAGCCCATGGATGGTATTAATTTCTGGGTG 2100
 Db 2044 AGCTGGCAAGAGGAGTGTGAGTCTTCAACAGCCCATGGATGGTATTAATTTCTGGGTG 2103
 Oy 2101 GGCATCCCGCAGGCGCTCAAGCATCGACATCCGCGAGCGGTATACAAGGGCTGATC 2160
 Db 2104 GGCATCCCGCAGGCGCTCAAGCATCGACATCCGCGAGCGGTATACAAGGGCTGATC 2163
 Oy 2161 GGGGATACAACTACTGCTGTGGAAGAACAGCCAAAGGCAAGTACTGCTCAACGGGGCAT 2220
 Db 2164 GGGGATACAACTACTGCTGTGGAAGAACAGCCAAAGGCAAGTACTGCTCAACGGGGCAT 2223
 Oy 2221 TTGCTGTGTGTGGGTGGAGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
 Db 2224 TTGCTGTGTGTGGGTGGAGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2283
 Oy 2281 GGCACGGGCAAGCGGTGGAGAGCTTGAGGCTTCCGGCCCATCTGAGACCGGCTGACC 2340
 Db 2284 GGCACGGGCAAGCGGTGGAGAGCTTGAGGCTTCCGGCCCATCTGAGACCGGCTGACC 2343
 Oy 2341 GTGAGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTATCTG 2400
 Db 2344 GTGAGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTATCTG 2403
 Oy 2401 CCCAAAGAGCTCGGAGAGCAAGTCTCATCCCAAGAACCCCGGGAGCCCTCTGTG 2460
 Db 2404 CCCAAAGAGCTCGGAGAGCAAGTCTCATCCCAAGAACCCCGGGAGCCCTCTGTG 2463
 Oy 2461 TTGCAACAACAGCTGTCTAGCTTCAACAGGTGGAGACCGCGGACGACAGGCCCCCT 2520
 Db 2464 TTGCAACAACAGCTGTCTAGCTTCAACAGGTGGAGACCGCGGACGACAGGCCCCCT 2523
 Oy 2521 GCACGCTGGGTGGGTGGAGAGTGGGGGCGTGTCCGCGAGACCTGGGCACTGGGCTGAG 2580
 Db 2524 GCACGCTGGGTGGGTGGAGAGTGGGGGCGTGTGTCCGCGAGACCTGGGCACTGGGCTGAG 2583

Oy 2581 AAGGGGCGGTGGAGTGTGGGGCTCGCGCGGACGACGAGGTCCTGCTGTGATGCA 2640
 Db 2584 AAGGGGCGGTGGAGTGTGGGGCTCGCGCGGACGACGAGGTCCTGCTGTGATGCA 2643
 Oy 2641 GCCCATGAGCCGCTGGAGACAAAGCTGGGGGAGCCCTGCCCCACTGGAGCTCAAC 2700
 Db 2644 GCCCATGAGCCGCTGGAGACAAAGCTGGGGGAGCCCTGCCCCACTGGAGCTCAAC 2703
 Oy 2701 GCTGTGTACCCCTGTCTCAAGACTGTGGGCGGGGATTTGAGAGGGCTCACTCAAGTGT 2760
 Db 2704 GCTGTGTACCCCTGTCTCAAGACTGTGGGCGGGGATTTGAGAGGGCTCACTCAAGTGT 2763
 Oy 2761 GTGGGCCACGAGGCGGCTGTGTGGCCCGGAGACAGTGAACCTTGACCGCAAGCCCCAG 2820
 Db 2764 GTGGGCCACGAGGCGGCTGTGTGGCCCGGAGACAGTGAACCTTGACCGCAAGCCCCAG 2823
 Oy 2821 GAGCTGAGCTTCTGCTGCTGAGAGCCGCTGTA 2853
 Db 2824 GAGCTGAGCTTCTGCTGCTGAGAGCCGCTGTA 2856

RESULT 4 US-10-009-332-2

: Sequence 2, Application US/10009332
 : GENERAL INFORMATION:
 : APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
 : APPLICANT: Kazusa DNA Research Institute
 : TITLE OR INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY
 : FILE REFERENCE: 067541
 : CURRENT APPLICATION NUMBER: US/10/009,332
 : PRIOR FILING DATE: 2001-12-10
 : PRIOR APPLICATION NUMBER: JPA Hei 11-321740
 : PRIOR FILING DATE: 1999-11-11
 : PRIOR APPLICATION NUMBER: JPA 2000-144020
 : NUMBER OF SEQ ID NOS: 35
 : SOFTWARE: Patentin version 3.1
 : SEQ ID NO 2
 : LENGTH: 2853
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-10-009-332-2

Query Match 99.8%; Score 2846.6; DB 38; Length 2853;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2849; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ATGCTTGTGGGTGGATCTTAACCTGGCTTTCGCGGGGCAACCGCTGAGGCTTGAG 60
 Db 1 ATGCTTGTGGGTGGATCTTAACCTGGCTTTCGCGGGGCAACCGCTGAGGCTTGAG 60
 Oy 61 CCAGAGCGGAGAGTGTGTTCCATTCGACTGAGACCGGCAATTACGGCCGCCCTAC 120
 Db 61 CCAGAGCGGAGAGTGTGTTCCATTCGACTGAGACCGGCAATTACGGCCGCCCTAC 120
 Oy 121 TACTGGCGGGTCCGAGAGACTCCGGGGATCAGGACATATTTTCAATACAGCATTT 180
 Db 121 TACTGGCGGGTCCGAGAGACTCCGGGGATCAGGACATATTTTCAATACAGCATTT 180
 Oy 181 CAGGAGAGCTTTTACTCACTGACGCGGATGCTAGTTCTTGGCTCCGCTTCC 240
 Db 181 CAGGAGAGCTTTTACTCACTGACGCGGATGCTAGTTCTTGGCTCCGCTTCC 240
 Oy 241 ACTGAGCATGTGGGCTCCCTCCAGAGGGCTCACCGGGGCTCTTCAGACTGAGAGC 300
 Db 241 ACTGAGCATGTGGGCTCCCTCCAGAGGGCTCACCGGGGCTCTTCAGACTGAGAGC 300
 Oy 301 TGTCTTATTTTGGGAGAGTGAAGCGGAGCGGAGTGGTGTGCTGTGAGCTGTGC 360
 Db 301 TGTCTTATTTTGGGAGAGTGAAGCGGAGCGGAGTGGTGTGCTGTGAGCTGTGC 360
 Oy 361 GGGGCGCTCCGCGGAGCCTTTGGCTAACGAGGCGCGAGTATGATTAAGCCCGCTGCC 420

|||||
Db 361 GGGGGCTCCGGAGCTTTGGCTACCGAGGCGCGGATGTCATTAGCCCGCTGCC 420
QY 421 AATGCTAGCG 480
Db 421 AATGCTAGCG 480
QY 481 GGTGTTCCGGGGCG 540
Db 481 GGTGTTCCGGGGCG 540
QY 541 AACCCCGCATCTTACGGGCGCTTGAGACCTTACAGCGCGCGCGCGCGCGCGCGCG 600
Db 541 AACCCCGCATCTTACGGGCGCTTGAGACCTTACAGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AGTCGTAGCG 660
Db 601 AGTCGTAGCG 660
QY 661 GAGAGCTGGGTGCGCGAGTCAATGTCAGTTCACCGCGCGCGCGCGCGCGCGCGCAT 720
Db 661 GAGAGCTGGGTGCGCGAGTCAATGTCAGTTCACCGCGCGCGCGCGCGCGCGCGCAT 720
QY 721 TATCTGCTGACGCTGCTGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 721 TATCTGCTGACGCTGCTGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 CCCATCAACATGCTTGGTGAAGTGTGCTTCTTAGAGATCGTACTCGGCGCGCGCGCG 840
Db 781 CCCATCAACATGCTTGGTGAAGTGTGCTTCTTAGAGATCGTACTCGGCGCGCGCGCG 840
QY 841 GTACCGGCAATGCG 900
Db 841 GTACCGGCAATGCG 900
QY 901 AAGTGTAGTGAACGACCGCGAGTACTGGGACACTGCGATCCCTTCACGAGCGGAC 960
Db 901 AAGTGTAGTGAACGACCGCGAGTACTGGGACACTGCGATCCCTTCACGAGCGGAC 960
QY 961 CTGTGTGAGCGCACCATCTGTGACACCTGGGCGATGGTGTGATGTCATGTGTGAC 1020
Db 961 CTGTGTGAGCGCACCATCTGTGACACCTGGGCGATGGTGTGATGTCATGTGTGAC 1020
QY 1021 CCCAAGAGAGTGTCTGTCTATGAGGACGATGGGTTCATCAGCGCTTCACGACTGCC 1080
Db 1021 CCCAAGAGAGTGTCTGTCTATGAGGACGATGGGTTCATCAGCGCTTCACGACTGCC 1080
QY 1081 CACGAGTGGGCG 1140
Db 1081 CACGAGTGGGCG 1140
QY 1141 TTTGGGAAGTCCGAGCGCAACCATGATGTCCCGGACCGCTTCATCCAGATCGACCGTGC 1200
Db 1141 TTTGGGAAGTCCGAGCGCAACCATGATGTCCCGGACCGCTTCATCCAGATCGACCGTGC 1200
QY 1201 AACCCCTGTGACGCTGCGAGTGTCCCATCATCAGGACTTCCTGGACGCGGCGCGGT 1260
Db 1201 AACCCCTGTGACGCTGCGAGTGTCCCATCATCAGGACTTCCTGGACGCGGCGCGGT 1260
QY 1261 GACTGCTCTGAGCAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db 1261 GACTGCTCTGAGCAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
QY 1321 AGCTACACCTTACGACGAGTGGAGTGGCTTTGGCGTGGGCTCCAGCCCTGTGCTT 1380
Db 1321 AGCTACACCTTACGACGAGTGGAGTGGCTTTGGCGTGGGCTCCAGCCCTGTGCTT 1380
QY 1381 TACATGCACTTACGACCAAGTGTGTGTCACCGGAGGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 TACATGCACTTACGACCAAGTGTGTGTCACCGGAGGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CAGACCG 1500
|||||

Db 1441 CAGACCG 1500
QY 1501 AAAGGGCGCTCGGTGGAGACACAACCTCAACAGCACAGGCTGGATGCTTCTGGGCG 1560
Db 1501 AAAGGGCGCTCGGTGGAGACACAACCTCAACAGCACAGGCTGGATGCTTCTGGGCG 1560
QY 1561 AAATGGATCCCTATGCGCCCTGCTCGCGACATGTGTGGGGCGGTGCGAGCTGGCGAG 1620
Db 1561 AAATGGATCCCTATGCGCCCTGCTCGCGACATGTGTGGGGCGGTGCGAGCTGGCGAG 1620
QY 1621 AGGAGTGCACCAACCCCGCTTCAAGCGGTACAAACGGGGGCAAGTACTGCGAGGAGT 1680
Db 1621 AGGAGTGCACCAACCCCGCTTCAAGCGGTACAAACGGGGGCAAGTACTGCGAGGAGT 1680
QY 1681 AAATACCGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCCTCGGAAAGAGCTTCGG 1740
Db 1681 AAATACCGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCCTCGGAAAGAGCTTCGG 1740
QY 1741 GAGGAGCAGTGTAGGCTTCAAGCGGTACAAACAGCACACCAACCGGCTCACTTCGCC 1800
Db 1741 GAGGAGCAGTGTAGGCTTCAAGCGGTACAAACAGCACACCAACCGGCTCACTTCGCC 1800
QY 1801 GTGGCATGGGTGCCCAAGTACTCGGGGTGTCTCCCGGGACAAAGTCAAGCTCATCTGC 1860
Db 1801 GTGGCATGGGTGCCCAAGTACTCGGGGTGTCTCCCGGGACAAAGTCAAGCTCATCTGC 1860
QY 1861 CGAGCAATGAGCAGTGTGCTTCTATGTGCTGGCACCAAGTGTGGAGCGGACGCTG 1920
Db 1861 CGAGCAATGAGCAGTGTGCTTCTATGTGCTGGCACCAAGTGTGGAGCGGACGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCGCTGTCTTCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCGCTGTCTTCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGAACCTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAATAG 2040
Db 1981 GGAACCTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAATAG 2040
QY 2041 AGCTCAAGAGAGTGTGACTGAGCTTCAACAGCCCATGCTGCTCAATTTCTGCTG 2100
Db 2041 AGCTCAAGAGAGTGTGACTGAGCTTCAACAGCCCATGCTGCTCAATTTCTGCTG 2100
QY 2101 GCATCCCGCGAGCGCGCTCAAGCATCGCATTCGCGACGCGGCTTACAAGGCTGATC 2160
Db 2101 GCATCCCGCGAGCGCGCTCAAGCATCGCATTCGCGACGCGGCTTACAAGGCTGATC 2160
QY 2161 GGGGATGACAACTACCTGGCTCTGAAGAACCAAGCAAGTACCTGCTCAAGCGGCA 2220
Db 2161 GGGGATGACAACTACCTGGCTCTGAAGAACCAAGCAAGTACCTGCTCAAGCGGCA 2220
QY 2221 TFCGTGGTGTGCGCGGTGAGCGGACCTGTGTGTAAGGCACTGTGCTGCGGTACAGC 2280
Db 2221 TFCGTGGTGTGCGCGGTGAGCGGACCTGTGTGTAAGGCACTGTGCTGCGGTACAGC 2280
QY 2281 GGCAGCGCACAGCGGTGGAGCGCTCGAGCTTCCCGGCGCATCTCGGAGCGCTGAC 2340
Db 2281 GGCAGCGCACAGCGGTGGAGCGCTCGAGCTTCCCGGCGCATCTCGGAGCGCTGAC 2340
QY 2341 GTGGAGTCTCTCGTGGGGAAGTACACCGCGCGGCTGCTGCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCGTGGGGAAGTACACCGCGCGGCTGCTGCTTCTATCTG 2400
QY 2401 CCCAAGAGCTTCGGAGGCAAGTCTCTCATCCAGGACCCCGGGGACCTCTGTC 2460
Db 2401 CCCAAGAGCTTCGGAGGCAAGTCTCTCATCCAGGACCCCGGGGACCTCTGTC 2460
QY 2461 TTGCAACAAGCGTCTCAGCTTCTTCAACAGGTGGAGCAGCGGAGCAGACAGGCGCC 2520
Db 2461 TTGCAACAAGCGTCTCAGCTTCTTCAACAGGTGGAGCAGCGGAGCAGACAGGCGCC 2520
QY 2521 GCACGCTGGGTGGGTGGGAGCTGGGGCGGTGCTCCGCGAGCTGGGCGAGTGGCGCTG 2580
Db 2521 GCACGCTGGGTGGGTGGGAGCTGGGGCGGTGCTCCGCGAGCTGGGCGAGTGGCGCTG 2580
|||||

QY 2581 AAGCGGGGGTGGAGTGTGGGGCTCCGGGGGACGACGAGTCCCTGCTGTGTATGCA 2640
Db 2581 AAGCGGGGGTGGAGTGTGGGGGCTCCGGGGGACGACGAGTCCCTGCTGTGTATGCA 2640
QY 2641 GCCATCGGGCCGTGGAGACACAAAGCTCGGGGGAGCCCTGCCACCTGGAGGCTCAGC 2700
Db 2641 GCCATCGGGCCGTGGAGACACAAAGCTCGGGGGAGCCCTGCCACCTGGAGGCTCAGC 2700
QY 2701 GCCTGTGTCACCTGCTCCAAAGAGCTCGGGCCGGGGATTTTCAGAGGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGTGTCACCTGCTCCAAAGAGCTCGGGCCGGGGATTTTCAGAGGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCCACGAGGCGCGGCTGTGGCCCGGACGAGTGGAACTTGCACGCAAGCCCCAG 2820
Db 2761 GTGGGCCACGAGGCGCGGCTGTGGCCCGGACGAGTGGAACTTGCACGCAAGCCCCAG 2820
QY 2821 GAGCTGAGCTTCTCGCTCTCGAGGCGGCTGCTGA 2853
Db 2821 GAGCTGAGCTTCTCGCTCTCGAGGCGGCTGCTGA 2853

RESULT 5

US-60-216-821-9
; Sequence 9, Application US/60216821
; GENERAL INFORMATION:
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Tang, Y. tom
; APPLICANT: Walsh, Roderick T.
; APPLICANT: Lal, Preeti
; APPLICANT: Khan, Farrah A.
; APPLICANT: Nguyen, Danniell B.
; APPLICANT: Tribouley, Catherine M.
; APPLICANT: Yue, Henry
; APPLICANT: Hafalia, April
; APPLICANT: Patterson, Chandra
; APPLICANT: Lu, Dyung Aina M.
; APPLICANT: Azimzai, Yaida
; APPLICANT: Gandhi, Ameena
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Yao, Monique G.
; APPLICANT: Lee, Ernestine A.
; APPLICANT: Lu, Yan
; APPLICANT: Ramkumar, Jaya
; APPLICANT: Kallick, Deborah A.
; APPLICANT: Au-Young, Janice
; APPLICANT: Xu, Yuming
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: FI-0146 P
; CURRENT APPLICATION NUMBER: US/60/216,821
; CURRENT FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473089CB1
US-60-216-821-9

Query Match 98.8%; Score 2817.6; DB 65; Length 2930;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2846; Conservative 0; Mismatches 4; Indels 6; Gaps 2;

QY 1 ATGCTTCTGCTGGGATCTTAACCTCGCTTCGCGGGGGAACCGCTGCGAGGCTCTGAG 60
Db 75 ATGCTTCTGCTGGGATCTTAACCTCGCTTCGCGGGGGAACCGCTGCGAGGCTCTGAG 134
QY 61 CCAGAGCGGGAGGTAGTGTGTTCCCATCCGAGTGGACCGGACATTAAAGCGCGCGGTAC 120
Db 135 CCAGAGCGGGAGGTAGTGTGTTCCCATCCGAGTGGACCGGACATTAAAGCGCGCGGTAC 194

QY 121 TACTGGGGGTCCCGAGGACTCCGGGGATCAGGAGTCAATTTTTCAGATCAGCAAT 180
Db 195 TACTGGGGGTCCCGAGGACTCCGGGGATCAGGAGTCAATTTTTCAGATCAGCAAT 254
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240
Db 255 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 314
QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGTCAACCGGGGCTCTTCAGAGCTCGGACGC 300
Db 315 ACTGAGCATCTGGGCTCCCGCTCCAGGGGTCAACCGGGGCTCTTCAGAGCTCGGACGC 374
QY 301 TGTCTTCTATTCTGGGAGCTGAACGCGGAGCTGCTGCTGCTGCTGCTGAGAGCTGTGC 360
Db 375 TGTCTTCTATTCTGGGAGCTGAACGCGGAGCTGCTGCTGCTGCTGCTGAGAGCTGTGC 434
QY 361 GGGGGGTCCCGGAGGCTTTGGCTACGAGGCGCGGATGTCATTTAGCCGCTGCC 420
Db 435 GGGGGGTCCCGGAGGCTTTGGCTACGAGGCGCGGATGTCATTTAGCCGCTGCC 494
QY 421 AATGCTAGCGCGCGCGCGGAGCAACAGCGAGGCGCACACCTTCTCCAGCGCGG 480
Db 495 AATGCTAGCGCGCGCGCGGAGCAACAGCGAGGCGCACACCTTCTCCAGCGCGG 554
QY 481 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCTTCGGGCTGG 540
Db 555 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCTTCGGGCTGG 614
QY 541 AACCCCGGCATCTACGGGCTCTGAGCCCTTACAGCGCGCGCGGGCTTCGGGGAG 600
Db 615 AACCCCGGCATCTACGGGCTCTGAGCCCTTACAGCGCGCGGGCTTCGGGGAG 674
QY 601 AGTCTAGCGCGCGAGTCTGGGCGCGCAAGCGTTTCGTGCTATCCCGGGTACGTG 660
Db 675 AGTCTAGCGCGCGAGTCTGGGCGCGCAAGCGTTTCGTGCTATCCCGGGTACGTG 734
QY 661 GAGAGCTGGTGGTGGCGGAGCTCAATGTCAGGTCAGCGCGCGGAGCTGCAACAT 720
Db 735 GAGAGCTGGTGGTGGCGGAGCTCAATGTCAGGTCAGCGCGCGGAGCTGCAACAT 794
QY 721 TATCTGCTGACGCTGCTGGCAACGGCGCGGAGCTTACCGCCATCCAGAGCTCTCAAC 780
Db 795 TATCTGCTGACGCTGCTGGCAACGGCGCGGAGCTTACCGCCATCCAGAGCTCTCAAC 854
QY 781 CCCATCAACATCGTGTGCTGAAGTGTGCTTCTTAGAGATCGTCACTCCGGGCCCCAAG 840
Db 855 CCCATCAACATCGTGTGCTGAAGTGTGCTTCTTAGAGATCGTCACTCCGGGCCCCAAG 914
QY 841 GTCACCGCAATCGCGGCTGACGCTGCGCAACTTCTGCTGCGCAGAGAGAGCTGAAC 900
Db 915 GTCACCGCAATCGCGGCTGACGCTGCGCAACTTCTGCTGCGCAGAGAGAGCTGAAC 974
QY 901 AAAGTGAGTGACAAGCACCCCGAGTACTGGGACACTGCGCATCTCTTCCAGAGGAGGAC 960
Db 975 AAAGTGAGTGACAAGCACCCCGAGTACTGGGACACTGCGCATCTCTTCCAGAGGAGGAC 1034
QY 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGGATGGTGTGGGTACCATGTGTGAC 1020
Db 1035 CTGTGTGGAGCCACCACTGTGACACCCCTGGGATGGTGTGGGTACCATGTGTGAC 1094
QY 1021 CCCAAGAGAGCTGCTCTGCTCAATTGAGGAGCTGGGCTTCCATGAGCCTTCCAGCTGCC 1080
Db 1095 CCCAAGAGAGCTGCTCTGCTCAATTGAGGAGCTGGGCTTCCATGAGCCTTCCAGCTGCC 1154
QY 1081 CAGAGCTGGGCGACGTGTTCACATGCCCATGACAAATGTGAAGCTCTGTGAGGAGGTG 1140
Db 1155 CAGAGCTGGGCGACGTGTTCACATGCCCATGACAAATGTGAAGCTCTGTGAGGAGGTG 1214
QY 1141 TTTGGGAAGCTCCGAGCAACACCATGATGTCCCCGACCTCATCCAGATCGACCGCTGCC 1200
Db 1215 TTTGGGAAGCTCCGAGCAACACCATGATGTCCCCGACCTCATCCAGATCGACCGCTGCC 1274

QY 1201 AACCCCTGTCAGCTGAGTCTGCCATCATCACCGACTTCCTTGAGACGGGACGGT 1260
DB 1275 AACCCCTGTCAGCTGAGTCTGCCATCATCACCGACTTCCTTGAGACGGGACGGT 1334
QY 1261 GACTGCTCTGAGACCAACCCAGACCCATCTCCCTGCCCGAGGATCTCCCGGGGCC 1320
DB 1335 GACTGCTCTGAGACCAACCCAGACCCATCTCCCTGCCCGAGGATCTCCCGGGGCC 1394
QY 1321 AGCTACACCTGAGCAGCAGTGGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCCT 1380
DB 1395 AGCTACACCTGAGCAGCAGTGGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCCT 1454
QY 1381 TACATGCACTACTGACCAAGCTTGTGTGACACCGGGAAGGCCAAGGACACATGGTGTG 1440
DB 1455 TACATGCACTACTGACCAAGCTTGTGTGACACCGGGAAGGCCAAGGACACATGGTGTG 1514
QY 1441 CAGACCGGCACTTCCTTGGGCGGATGGCAGCAGCTGTGGCGAGGGCAAGCTTCGCTC 1500
DB 1515 CAGACCGGCACTTCCTTGGGCGGATGGCAGCAGCTGTGGCGAGGGCAAGCTTCGCTC 1574
QY 1501 AAGGGGCTGCTGGAGAGACACAACCTCAACAGACACAGGGTGGATGGTTCCTGGGCC 1560
DB 1575 AAGGGGCTGCTGGAGAGACACAACCTCAACAGACACAGGGTGGATGGTTCCTGGGCC 1634
QY 1561 AATAGGATCCCTATGAGCCCTGCTGCGGCACATGTGTGGGGGGCTGACAGCTGGCCAGG 1620
DB 1635 AATAGGATCCCTATGAGCCCTGCTGCGGCACATGTGTGGGGGGCTGACAGCTGGCCAGG 1694
QY 1621 AGGAGTGCACCAACCCACCCCTGCCAAGCGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
DB 1695 AGGAGTGCACCAACCCACCCCTGCCAAGCGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1754
QY 1681 AATACCGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCTCAGCTCCGGAAAGAGCTCCGG 1740
DB 1755 AATACCGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCTCAGCTCCGGAAAGAGCTCCGG 1814
QY 1741 GAGGACAGTGTGAGGCTTTCAACGGCTTACACACAGCACCAACCGGCTCACCTCGCC 1800
DB 1815 GAGGACAGTGTGAGGCTTTCAACGGCTTACACACAGCACCAACCGGCTCACCTCGCC 1874
QY 1801 GTGGCATGGTCCCAAGTACTCCGGCTGTCTCCCGGGGACAGTGCACGCTCATCTGC 1860
DB 1875 GTGGCATGGTCCCAAGTACTCCGGCTGTCTCCCGGGGACAGTGCACGCTCATCTGC 1934
QY 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGCACCCAA---GGTGGTGGAGGGCAGC 1917
DB 1935 CGAGCAATGGCACTGGCTTCTATGTGTGGCACCCAAAGGTGTGTGGAGCGCAG 1994
QY 1918 CTGTGCTCTGCTGACTCCACCTCGCTGTGTGCAAGGCAAGTGCATCAAGCTGGCTGT 1977
DB 1995 CTGTGCTCTGCTGACTCCACCTCGCTGTGTGCAAGGCAAGTGCATCAAGGCTGGCTGT 2054
QY 1978 GATGGAACTGGCTGCCAAGAGATTCGACAGTGTGGGGTGTGGGGGAGACAA 2037
DB 2055 GATGGAACTGGCTGCCAAGAGATTCGACAGTGTGGGGTGTGGGGGAGACAA 2114
QY 2038 AAGAGTGCAGAAGGTGACTGGACTTTCACCAAGCCCATTCATGGCTTCAATTTCTGT 2097
DB 2115 AAGAGTGCAGAAGGTGACTGGACTTTCACCAAGCCCATTCATGGCTTCAATTTCTGT 2174
QY 2098 GTGGCATCCCGCAGGCGCTCAAGCATCGACATCCCGCGGTTACAAAGGGCTG 2157
DB 2175 GTGGCATCCCGCAGGCGCTCAAGCATCGACATCCCGCGGTTACAAAGGGCTG 2234
QY 2158 ATCGGGGATGACAACCTACCTGGCTGTGAAGAAGCAGCAGGCAAGTACTGCTCAACGGG 2217
DB 2235 ATCGGGGATGACAACCTACCTGGCTGTGAAGAAGCAGCAGGCAAGTACTGCTCAACGGG 2294
QY 2218 CATTTGTTGTTGCTGGCGGTGAGCGGGACCTGTGTGTGAAGGGCAGTCTGCTGGCGTAC 2277
DB 2295 CATTTGTTGTTGCTGGCGGTGAGCGGGACCTGGTGTGTGAAGGGCAGTCTGCTGGCGTAC 2354
QY 2278 AGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAGGCTTCCTCCGCGCCCATCTCCGAGCGCTG 2337

DB 2355 AGCGGCACGGGCACAGCGGTGGAGAGCCTGCAGGCTTCCTCCGCGCCCATCTCTGGAGCGCTG 2414
QY 2338 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTCTAT 2397
DB 2415 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2474
QY 2398 CTGCCAAAGACCTTCGGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGG---GGACCC 2454
DB 2475 CTGCCAAAGACCTTCGGGAGGACAAAGTCTCTCATCCCGCGCACCCCGGGGAGAGACC 2534
QY 2455 TCTGTCTTGCACACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCGCGGACGACAGG 2514
DB 2535 TCTGTCTTGCACACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCGCGGACGACAGG 2594
QY 2515 CCCCTGCAAGCTGGTGGCTGGCAGCTGGGGGCGCTGCTCCGCGAGCTGGCGGAGTGGC 2574
DB 2595 CCCCTGCAAGCTGGTGGCTGGCAGCTGGGGGCGCTGCTCCGCGAGCTGGCGCACTGGC 2654
QY 2575 CTGCAGAGCGGGCGGTGGACTGTGGGGCTCCGCGGGCAGCGACGCTCCCTGCTGT 2634
DB 2655 CTGCAGAGCGGGCGGTGGACTGTGGGGGCTCCGCGGGCAGCGACGCTCCCTGCTGT 2714
QY 2635 GATGACGCCCATTCGGCCGCTGGAGACACAAGCCTTCGGGGAGCCCTGCCACCTGGGAG 2694
DB 2715 GATGACGCCCATTCGGCCGCTGGAGACACAAGCCTTCGGGGAGCCCTGCCACCTGGGAG 2774
QY 2695 CTCAGGCTGTGCTACCTGTCTCCAAAGACTTCGGGGGAGCCCTGCCACCTGGGAG 2754
DB 2775 CTCAGGCTGTGCTACCTGTCTCCAAAGACTTCGGGGGAGCCCTGCCACCTGGGAG 2834
QY 2755 AAGTGTGTGGGCGCAGGCGGCTGTGTGGCGGGGAGCTGGCGCGGGGATTCAGAGCGCTCACTC 2814
DB 2835 AAGTGTGTGGGCGCAGGCGGCTGTGTGGCGGGGAGCTGGCGCGGGGATTCAGAGCGCTCACTC 2894
QY 2815 CCCAGGAGCTGGACTTCTGCTCTGAGCGCGTGC 2850
DB 2895 CCCAGGAGCTGGACTTCTGCTCTGAGCGCGTGC 2930

RESULT 6

US-10-093-463-27

; Sequence 27, Application US/10093463

; GENERAL INFORMATION:

; APPLICANT: Padigar, Muralidhara

; APPLICANT: Shenoy, Suresh

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Gusev, Vladimir

; APPLICANT: Pochart, Pascal

; APPLICANT: Zhong, Mei

; APPLICANT: Rastelli, Luca

; APPLICANT: Mezes, Peter

; APPLICANT: Smithson, Glenn

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gerlach, Valerie

; APPLICANT: Casman, Stacie

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Tchernev, Velizar

; APPLICANT: Gangolli, Esha

; APPLICANT: Vernet, Corine

; APPLICANT: Pena, Carol

; APPLICANT: Burgess, Catherine

; APPLICANT: Liu, Xiaohong

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gorman, Linda

; APPLICANT: Spaderna, Steven

; APPLICANT: Voss, Edward

; APPLICANT: Malyankar, Uriel

; APPLICANT: Anderson, David

; APPLICANT: Patturajan, Meera

; APPLICANT: Miller, Charles

APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids, and Methods of Use.
FILE REFERENCE: 21402-290A (Cura 590AT)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 27
LENGTH: 2804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (25)..(2797)
US-10-093-463-27

Query Match 91.88; Score 2619,2; DB 39; Length 2804;

Best Local Similarity 96.28; Pred No. 0; Mismatches 13; Indels 96; Gaps 3;

Matches 2753; Conservative 0;

1 ATGCTTCTGCTGGGATCTAACTGCTTCCCGGGGAAACCGCTGGAGGCTCTGAG 60

25 ATGCTTCTGCTGGGATCTAACTGCTTCCCGGGGAAACCGCTGGAGGCTCTGAG 84

61 CCAGAGGGGAGGTAGTCTCCATCGGACTGGACCGGACATTAACGGCCCGCTAC 120

85 CCAGAGGGGAGGTAGTCTCCATCGGACTGGACCGGACATTAACGGCCCGCTAC 144

121 TACTGGGGGTCCTGGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCACAGCATTT 180

145 TACTGGGGGTCCTGGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCACAGCATTT 204

181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGTCCCGGCTCTCC 240

205 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGTCCCGGCTCTCC 264

241 ACTGAGCATCTGGCGTCCCTCCAGGGGTACCGGGGGCTCTTCAGACCTTCGGAGCG 300
265 ACTGAGCATCTGGCGTCCCTCCAGGGGTACCGGGGGCTCTTCAGACCTTCGGAGCG 324
301 TGCCTTCTATCTGGGACGTGAAGCCGAGCCGACGCTGCTGCTGTGAGGCTGTGC 360
325 TGCCTTCTATCTGGGACGTGAAGCCGAGCCGACGCTGCTGCTGTGAGGCTGTGC 384
361 GGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCCGAGTATGTCAATAGCCCTGCC 420
385 GGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCCGAGTATGTCAATAGCCCTGCC 444
421 AATGCTAGCGCGCGGCGGCGAGCGAACAAGCAGGGCGGACACCTTCTCCAGCGCG 480
445 AATGCTAGCGCGCGGCGGCGAGCGAACAAGCAGGGCGGACACCTTCTCCAGCGCG 504
481 GGTGTTCCGGGCGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCCGGGTG 540
505 GGTGTTCCGGGCGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCCGGGTG 564
541 AACCCCGCATCTACGGGCGCTTGAAGCCCTTACAAGCGCGGGGGGCTTCGGGAG 600
565 AACCCCGCATCTACGGGCGCTTGAAGCCCTTACAAGCGCGGGGGGCTTCGGGAG 624
601 AGTCTAGCGCGCGGCGGCTTGGGCGGCGGAGCGTTCGTGCTATCCCGGCTACGTG 660
625 AGTCTAGCGCGCGGCGGCTTGGGCGGCGGAGCGTTCGTGCTATCCCGGCTACGTG 684
661 GAGAGCTGGTGGTGGCGGAGGCTCAATGGTTCCACGGCGGCGGAGCTTGAACAT 720
685 GAGAGCTGGTGGTGGCGGAGGCTCAATGGTTCCACGGCGGCGGAGCTTGAACAT 744
721 TATCTGCTGAGCGTCTCTGGCAACGGCGGCGGAGCTTACCGGCGGAGCTTCAAC 780
745 TATCTGCTGAGCGTCTCTGGCAACGGCGGCGGAGCTTACCGGCGGAGCTTCAAC 804
781 CCATCAACATCTGTTGTTCAAGTGTCTTCTTAGAGATCGTGACTCCGGGCGGAG 840
805 CCATCAACATCTGTTGTTGTTCAAGTGTCTTCTTAGAGATCGTGACTCCGGGCGGAG 864
841 GTACCGGCAATCGCGGCGCTGACGTGGCGCAACTTCTGTGCTGGCAAGAGCTGAAC 900
865 GTACCGGCAATCGCGGCGCTGACGTGGCGCAACTTCTGTGCTGGCAAGAGCTGAAC 924
901 AAAGTGAGTGACAGCAACCGGCGGAGCTTGGGAGCTGCGGAGCTTCCATCCAGGAG 960
925 AAAGTGAGTGACAGCAACCGGCGGAGCTTGGGAGCTGCGGAGCTTCCATCCAGGAG 984
961 CTGTGTGGAGCCACCGCTGTGACACCGCTTGGGAGCTGCTGATGGGTACCAATGTG 1020
985 CTGTGTGGAGCCACCGCTGTGACACCGCTTGGGAGCTGCTGATGGGTACCAATGTG 1044
1021 CCCAAGAGAGTGTCTGTCTCAATGAGGAGCTGGGCTTCCATAGCCTTCCACTGCC 1080
1045 CCCAAGAGAGTGTCTGTCTCAATGAGGAGCTGGGCTTCCATAGCCTTCCACTGCC 1104
1081 CAGGAGCTGGGCGGAGCTGTTCAACATGCCCATGACCAATGTGAAAGTCTGTGAGGAG 1140
1105 CAGGAGCTGGGCGGAGCTGTTCAACATGCCCATGACCAATGTGAAAGTCTGTGAGGAG 1164
1141 TTTGGAAAGCTCCGAGCGCAACCAACATGATGTCCCGGAGCCTCATCCAGATCGACCG 1200
1165 TTTGGAAAGCTCCGAGCGCAACCAACATGATGTCCCGGAGCCTCATCCAGATCGACCG 1224
1201 AACCCCTGTGTAGCTGTGAGTGTGCTGCTATCATCCGAGCTTCTTGGAGAGCGGAG 1260
1225 AACCCCTGTGTAGCTGTGAGTGTGCTGCTATCATCCGAGCTTCTTGGAGAGCGGAG 1284
1261 GACTGCTCTTGGAGCAACCGGAGCGGAGCTTCCCTGCGGAGGATCTGCGGGCGGCG 1320
1285 GACTGCTCTTGGAGCAACCGGAGCGGAGCTTCCCTGCGGAGGATCTGCGGGCGGCG 1344
1321 AGCTACACCTTGAGCGGAGCTGCGAGCTGGCTTTTGGGCTGGGCTTCCAAAGCCCTG 1380

Db 1345 AGTACACCTGAGCAGAGTGCAGGCTGCTTTTGGCGTGGGCTCCAAAGCCCTGCTCT 1404
QY 1381 TACATCAGTACTGACCAAGCTGTGTCACCGGGAAGGCCAAGGACATGCTGTC 1440
Db 1405 TACATCAGTACTGACCAAGCTGTGTCACCGGGAAGGCCAAGGACATGCTGTC 1464
QY 1441 CAGACCGGCACTTCCCTGGGCGGATGGCAGCAGCTGTGCGAGGGCAAGCTGTGCTC 1500
Db 1465 CAGACCGGCACTTCCCTGGGCGGATGGCAGCAGCTGTGCGAGGGCAAGCTGTGCTC 1524
QY 1501 AAGGGGCTGCGTGAGAGACACACCTCAACAGCAC-----AGGGTGAATGGT 1551
Db 1525 AAGGGGCTGCGTGAGAGACACACCTCAACAGCAC-----AGGGTGAATGGT 1584
QY 1552 TCCTGGGCAAAATGGGATCCCTATGCCCCCTGCTCGCCACATGTGTGGGGGCTGCAG 1611
Db 1585 TCCTGGGCAAAATGGGATCCCTATGCCCCCTGCTCGCCACATGTGTGGGGGCTGCAG 1644
QY 1612 CTGGCCAGGAGCAGTGCACCAACCCACCCTGCTCCACAGGGGGCAAGTACTGAGGGA 1671
Db 1645 CTGGCCAGGAGCAGTGCACCAACCCACCCTGCTCCACAGGGGGCAAGTACTGAGGGA 1704
QY 1672 GTGAGGTGAATACCGATCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCCCTCGGAAAG 1731
Db 1705 GTGAGGTGAATACCGATCTGCAATCTGAGAGCCCTGCCCCAG-----CTCCGGAAAG 1758
QY 1732 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCACAGGCTTACACACAGCAGCAGCAGCGGCTC 1791
Db 1759 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCACAGGCTTTCACACAGCAGCAGCAGCGGCTC 1818
QY 1792 ACTCTCGCGTGCATGGGTGCCCAAGTACTCCGCGCTGTCTCCCGGAGCAAGTGCAG 1851
Db 1819 ACTCTCGCGTGCATGGGTGCCCAAGTACTCCGCGCTGTCTCCCGGAGCAAGTGCAG 1878
QY 1852 CTCATCTGCGAGCCCAATGGCAGTGGCTACTTATGTGCTGGCACCCCAAGTGGTGAC 1911
Db 1879 CTCATCTGCGAGCCCAATGGCAGTGGCTACTTATGTGCTGGCACCCCAAGTGGTGAC 1938
QY 1912 GGCACGCTGTCTCTCTGACTCCACCTCCGCTGTGTCAGAGCAAGTGCATCAAGCT 1971
Db 1939 GGCACGCTGTCTCTCTGACTCCACCTCCGCTGTGTCAGAGCAAGTGCATCAAGCT 1998
QY 1972 GCGTGTGATGGGAACCTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGGGGA 2031
Db 1999 GCGTGTGATGGGAACCTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGGGGA 2058
QY 2032 GACAATAAGAGCTGCAAGAGTGGTCTTCAAGAGATTCGACAAAGTGTGGGTGTGGGGA 2091
Db 2059 GACAATAAGAGCTGCAAGAGTGGTCTTCAAGAGATTCGACAAAGTGTGGGTGTGGGGA 2118
QY 2092 TTCGTGTGAGGCAATCCCGCAGCGGCTCAAGCATCGACATCCGCGAGCGGGTTACAAA 2151
Db 2119 TTCGTGTGAGGCAATCCCGCAGCGGCTCAAGCATCGACATCCGCGAGCGGGTTACAAA 2178
QY 2152 GGCCTGATCGGGATGACAACTACCTGGCTCTGAAGACAGCAAGCAAGTACCTGCTC 2211
Db 2179 GGCCTGATCGGGATGACAACTACCTGGCTCTGAAGACAGCAAGCAAGTACCTGCTC 2238
QY 2212 AACGGGCAATTCGTGTGTGCGGCTGGAGCGGAGCTGTGCTGAAGGGAGTGTGCTG 2271
Db 2239 AACGGGCAATTCGTGTGTGCGGCTGGAGCGGAGCTGTGCTGAAGGGAGTGTGCTG 2298
QY 2272 CGGTACAGCGGCACGGGACAGCGGTGGAGAGCTTGCAGGCTTCCCGGCCCATCTCGGAG 2331
Db 2299 CGGTACAGCGGCACGGGACAGCGGTGGAGAGCTTGCAGGCTTCCCGGCCCATCTCGGAG 2358
QY 2332 CCCTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGTCGGCTACATCC 2391
Db 2359 CCCTGACCGTGGAGTCTCTCTCCGTGGGGAAGATGACACCGCCCGGTCGGCTACATCC 2418
QY 2392 TTCTATCTGCCCCAAGAGCTCGGGAGGACAAAGTCTCTCATPCCCAAGACCCCGCGGGA 2451

Db 2419 TTCTATCTGCCCAAGAGCCTCGGAGGACAAAGTCTCTCTCATC----- 2461
QY 2452 CCCTCTGTCTTGACAAACAGCGTCTCAGCTCTCCAAACAGGTGGAGCAGCGGACGAC 2511
Db 2462 ----- 2461
QY 2512 AGCCGCCCTGACAGCTGGGTGGCTGGCAGCTGGGGGGCGGTGCTCCGGAGAGTGGCGGAGT 2571
Db 2462 -----CCCGCGCAGCTGGGTGGCTGGCAGCTGGGGGGCGGTGCTCCGGAGAGTGGCGGAGT 2517
QY 2572 GGCCTGCAAAAGCGGGCGGTGACTGTGGGGCTCCGCGGGGACGCGCAGCGTCCCTGCC 2631
Db 2518 GGCCTGCAAAAGCGGGCGGTGACTGTGGGGCTCCGCGGGGACGCGCAGCGTCCCTGCC 2577
QY 2632 TGTGATGACAGCCATCGGCCCTGGAGACACAAGCCTCGGGGAGCCCTGCCCCACCTGG 2691
Db 2578 TGTGATGACAGCCATCGGCCCTGGAGACACAAGCCTCGGGGAGCCCTGCCCCACCTGG 2637
QY 2692 GAGCTCAGCGCTGTGTCACCTGCTCAAGAGCTCGCGCGGGGATTTACAGAGCGGTCA 2751
Db 2638 GAGCTCAGCGCTGTGTCACCTGCTCAAGAGCTCGCGCGGGGATTTACAGAGCGGTCA 2697
QY 2752 CTCAGTGTGTGGGCCACGAGGCGGCTGTGGGCCCGGGACCAAGTGCACCTTGCACGCG 2811
Db 2698 CTCAGTGTGTGGGCCACGAGGCGGCTGTGGGCCCGGGACCAAGTGCACCTTGCACGCG 2757
QY 2812 AAGCCCCAGGAGCTGGACTTCTGCGTCTGAGGCGGCTGCTGA 2853
Db 2758 AAGCCCCAGGAGCTGGACTTCTGCGTCTGAGGCGGCTGCTGA 2799

RESULT 7
US-60-360-207-12354
; Sequence 12354, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 12354
; LENGTH: 5714
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-12354

Query Match 80.6%; Score 2299.4; DB 80; Length 5714;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 2507; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCATCTCAACCCCTGGCTTTTCGCCGGGCGAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGCATCTCAATCTCTGGCTCTCGCTGGCGACCTGCTGGCAGCTCCGAG 60
QY 61 CCAGACGGGAGGTAGTCTGTTCCCTCCACTGGACCCGGACATTAAGCGCGCGGCTAC 120
Db 61 CCAGAGTGGGAGGTGCTGTTCCATTCACACGGGACCGGACATCAATGGCGCGGCTAC 120
QY 121 TACTGCGGGGCTCCCGAGGAGCTCCGGGGATCAGGAGCTCATTTTTCAGATCACACATTT 180
Db 121 TACAGAGGGGTACGAGGAGCTCCGGGATCAGGGTCTCATTTTTCAGATCACACATTT 180
QY 181 CAGAGAGCTTTTACCTACCTGAGCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGCAGAGCTTTTATCTACCTGAGCGCGAGATGCCAGTTCCTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGCGTCCCGCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTGCGAGGC 300
Db 241 ACTGAGTATCTAGTGTCCCGCTGACAGAGGCTCAGTGGCAGCTCTCTAGACCTGCGAGGC 300
QY 301 TGGTCTTATCTCTGGGAGCTGAAACCGCGAGCGGAGCTCGTTCGCTGTGTGAGCTGTGC 360

Db 301 TGCTTCTATCCGGGTATGTGAACGGGAGCCAGACTCCTTTGCTGCTGTGAGCCATATGC 360
QY 361 GGGGGGCTCGGGAGACCTTTGGCTACCGAGGGCGGAGTATGTATTAAGCCCGCTGCCC 420
Db 361 GGGGGTCTCGCGGAGCCTTTGGCTACCGAGGTGCGGAGTATGTATTAAGCCCGCTGCCC 420
QY 421 AATGCTAGCGCGCGCGGCGAGCGCAACAGCAGCGGCGCACACCTTCTCCAGCGCCGG 480
Db 421 AACACGACGCGCGAGAGCGGAGCTCAGCGGCGGCGCACACCTTCTCCAGCGCCGG 480
QY 481 GGTGTTCCGGCGGCGCTTCCGAGAGCCCACTCTCGTGGGGGTGCGCTCGGGCTGG 540
Db 481 GGTGCTCTAGGGCTTCCGAGAGCCCACTCTCGTGGGGGTGCGCTCGGGCTGG 540
QY 541 AACCGCCCATCTACGGGCCCTGGAGCCCTTACAGCGCGGCGGCGGCGCTCGGGGAG 600
Db 541 AACCGCCCATCTAGGGCTTGGAGCCCTTAAAGCCAGCGGCGGCGGCGGCGGAG 600
QY 601 AGTCGTAGCGCGCGAGCTTGGCGCGCGCAAGCGTTTCGTCTATCCCGGGGTACGTTG 660
Db 601 AGCCACAACCGCGAGGCTTGGCGCGCGCAAGCGCTTCTGTCTATACACGGTACGTTG 660
QY 661 GAGACGCTGTGTGGCGGAGAGTCAATGGTCAAGTTTCCAGCGCGGCGGCGGAGACAT 720
Db 661 GAGACATGGTGGTGGCGGAGAGTCAATGGTCAAGTTTCCAGCGCGGCGGAGATTTGSAACAT 720
QY 721 TATCTGCTGAGCTGTGGCAAGCGGCGGCGGAGCTCTACCGCCATCCAGCATCTCAAC 780
Db 721 TATCTGCTGAGCTGTGGCGGCGGCGGCGGAGCTCTACCGCCATCCAGCATCTCAAC 780
QY 781 CCCATCAACATCGTGTGGTCAAGGTGCTCTTTAGAGATCGTGAATCGGCGGCGGAG 840
Db 781 CCTATCAACATCGTGTGGTCAAGGTGCTCTTTAGAGATCGTGAATCGGCGGCGGAG 840
QY 841 GTCACCGCAATGGCGCCCTGAGCTCGCAACTTCTGCTGCTGGCAGAGAGCTGAAC 900
Db 841 GTCACAGCAACGGCGGCGGCGGAGCTCTGCGCAACTTCTGCTGCTGGCAGAGAGCTGAAC 900
QY 901 AAAGTGAGTGAACAGCAGCCCGAGTACTGGGACACTGCCATCTCTTACACGAGCAGGAC 960
Db 901 AAAGTGAGCAGTAAAGCAGCCCTGAGTACTGGGACACAGCCATCTCTTACACGAGCAGGAC 960
QY 961 CTGTGTGAGGACCACTGTGTGACACCCCTGGGATGCTGATGGGTACCATGTGAC 1020
Db 961 CTATGCGGGGTACCACTGTGTGACACCTGGGATGCTGATGGGTACCATGTGAC 1020
QY 1021 CCCAAGAGAACTGCTGTCTATTGAGCAGATGGGCTTCCATCAGCCTTCAACACTGCC 1080
Db 1021 CCCAAGAGAACTGCTGTCTATTGAGCAGATGGGCTTCCATCAGCCTTCAACACTGCC 1080
QY 1081 CACGAGCTGGCCACGCTTCAACATGCCCATGCCCATGCAATGTGAAGTCTGTGAGGAGTG 1140
Db 1081 CATGAGCTGGCCATGTGTTCAACATGCCCATGCCCATGCAACGTGAAGTGTGTGAGGAGTG 1140
QY 1141 TTTGGAGAGCTCCGAGCCCAACCATGATGTCCCGGAGCCCTCATCCAGATGACCGTGC 1200
Db 1141 TTTGGAGAGCTCAGAGCCCAACCATGATGTCCCGGAGCCCTCATCCAGATGACCGTGC 1200
QY 1201 AACCCCTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 AACCCCTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 GACTGCTCTGAGCCAGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1320
Db 1261 GACTGCTCTCTGAGCCAGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1320
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1380
Db 1321 AGCTACAGCTTGGAGCCAGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1380
QY 1381 TACATGCACTGCTGACCAAGAGTGTGTGTCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1381 TATATGCACTGCTGACAAAGTGTGTGTCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

QY 1441 CAGACCGCCACTTCCCTTGGCGGATGGACACAGCTGTGGCGAGGCAAGCTTGGCTC 1500
Db 1441 CAGACTGCCACTTCCCTTGGCGAGATGGACACAGCTGTGGGAAGCAAGTTTGGCTC 1500
QY 1501 AAAGGGGCTCGGTGGAGAGACAACTCAACAGACAGGAGGATGGTCTTGGGCG 1560
Db 1501 AAGGAGGCTCGGTGGAGAGACAACTCAACAGTACCGGTGGAGGCTCTTGGGCG 1560
QY 1561 AAATGGGATCCTATGCGCCTGCTCGGCACATGTGGTGGGGGCTGCAGCTGGCCAGG 1620
Db 1561 AAGTGGGAGCCTTACGTTCTGCTCCGACCTTCCGCTGGGGGCTGCAGCTGGCCCG 1620
QY 1621 AGGAGTGCACCAACCCCTTCCCAAGGGGCAAGTACTCGAGGAGGAGTGGGCTG 1680
Db 1621 AGGAGTGCACCAACCCCTTACCCCTGCCAAGGTGGAAAACTCGAGGAGTAAAGT 1680
QY 1681 AAATACCATCTGCAATCTGGAGCCTTCCAGCGCTTACACACAGCAGCAGCTTCCCG 1740
Db 1681 AAATACCATCTTGCACCTTGGAGCCTTCCAGCGCTTACACACAGCAGCAGCTTCCCG 1740
QY 1741 GAGGAGAGTGTGAGGCTTTCAGCGGCTTACACACAGCAGCAGCAGCTTCTCCCG 1800
Db 1741 GAGGAGCAATGTGAGCTTTCAGTGGCTTACACACAGCAGCAGCAGCTTCTAGCT 1800
QY 1801 GTGGCATGGTGGCGGCAAGTACTCCCGGCTGTCTCCCGGAGCAAGTGCAGCTCATCTGC 1860
Db 1801 GTGGCATGGTGGCGGCTGTACACAGTGTACACAGTGTACAGCTTCTAGCTCATCTGC 1860
QY 1861 CGAGCAATGGAGCTGCTTCTATGCTGGCAGCAGGCTGGTGGAGCGGCGGCTG 1920
Db 1861 CGAGCAATGGAGCTGCTTCTATGCTGGCAGCAGGCTGGTGGAGCGGCGGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTTCCGCTGTGTCCAGGCAAGTGCATCAAGCTTCTGCTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTTCCGCTGTGTCCAGGCAAGTGCATCAAGCTTCTGCTGAT 1980
QY 1981 GGAACCTGGCTTCCAGAGAGATTCGACAGTGTGGGCTGTGGGGGAGACAATAAG 2040
Db 1981 GGAATCTGGCTTCCAGAGAGATTCGACAGTGTGGGCTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTCAAGAAAGTGAAGTCTTCCAGGCAAGCTTCCAGGCTTCTGCTTCTGCTGAT 2100
Db 2041 AGCTCAAGAAAGTGAAGTCTTCCAGGCAAGCTTCCAGGCTTCTGCTTCTGCTGAT 2100
QY 2101 GCAATCCCGAGGCGCTTCAAGCATCGACATCCCGAGCGGCTTCAAGGCGCTGATC 2160
Db 2101 GCAATCCCTGCGGGCGCTTCCAGCATTGTATCCGCTGAGCGGCTTCAAGGCGCTGAT 2160
QY 2161 GGGATGCAACTACTGCTGCTGAAGAACAGCAAGCAAGTACTGCTCAACCGGCT 2220
Db 2161 GGGATGCAACTACTGCTGCTGAAGAACAGCAAGCAAGTACTGCTCAACCGGCT 2220
QY 2221 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2221 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY 2281 GGCAGCGGACAGCGGTGAGAGCCTGAGCTTCCCGGCGCTTCCGAGCGGCTGAGC 2340
Db 2281 GGCAGCGGACAGCGGTGAGAGCCTGAGCTTCCCGGCGCTTCCGAGCGGCTGAGC 2340
QY 2341 GTGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2341 GTGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
QY 2401 CCCAAGAGCCTGCGGAGGAGCAAGTCTCTATCCAGAGCCTCCCGGGGAGCCTTCTGTC 2460
Db 2401 CCCAAGAGCCTGCGGAGGAGCAAGTCTCTATCCAGAGCCTCCCGGGGAGCCTTCTGTC 2460
QY 2461 TTGCAACAAGCCTTCTGAGCCTTCCAGAGGCTGGAGCAGCGGAGCAGAGGCGGCT 2520
Db 2461 TTGCGCAACAGTGTCTGAGCCTTCCAGAGGCTGGAGCAGCGGAGCAGAGCCTCT 2520


```
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGCTGCTCCGAGAGCTGGCGAGTGGCGCTGCGAG 2580
Db 2521 GCGCGCTGGTGGCAGCAGCTGGGACCTTGCTCAGTGAAGTGGCGAGTGGCGCTACAG 2580
QY 2581 AAGCGGGGGTGGAGTGTGGGGGTCGCCGGGCGACGACGGTCCCTGCTGCTGTGATGCA 2640
Db 2581 AAGCGAGCAGTGGAGTGTGGGGGTCGCCGGGCGACGACGGTCCCTGCTGTGATGTA 2640
QY 2641 GCCCATCGCGCGTGGAGACACAGCTTGGCGGGGAGCGCTGCCCGACCTGGAGCTCAGC 2700
Db 2641 GACCATCGCGCATGGAGAGCGAGCTGTGGGGAACCCCTGCCAACTTGGAGCTCGGC 2700
QY 2701 GCTTGGTCAACCTGTCTCAAGAGCTGGCGCGGGGATTTACAGAGCGCTCACTCAAGTGT 2760
Db 2701 AACTGGTGGCCCTGTCTTAAAGCTGTGGCCGTGGATTTAAGAGCGCTCCGCTCAAGTGT 2760
QY 2761 GTGGGCCACGAGGCGCGCTGCTGGCCGGGACCAAGTCACTTGCACCGCAAGCCCGCAG 2820
Db 2761 GTGGGCCACGAGGCGAGCTGCTGGCTGGGACCAAGTGTGACTTACGCGCAAGCCCGCAG 2820
QY 2821 GAGCTGGAGTCTCGCTGAGCGCTGCTGCTGA 2853
Db 2821 GAATTAGACTTCTGCTGCTTGAGACCCCTGCTGA 2853

RESULT 8
US-10-163-316-3
; Sequence 3, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163, 316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-10-163-316-3

Query Match 80.5%; Score 2298; DB 41; Length 2469;
Best Local Similarity 97.0%; Pred. No 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

QY 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCCCGGGCGAACCCTGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCCCGGGCGAACCCTGAGGCTCTGAG 60
QY 61 CCAGAGGGGAGTGTGCTTCCATCCGACTGGACCCGACATTAACCGCGCGCGCTAC 120
Db 61 CCAGAGGGGAGTGTGCTTCCATCCGACTGGACCCGACATTAACCGCGCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGCATTT 180
Db 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGCATTT 180
QY 181 CAGGAGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCC 240
Db 181 CAGGAGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCC 240
QY 241 ACTGACACTTGGCGCTCCCGCTCCAGGGCTCACCAGGGCTCTTTCAGACCTGGAGC 300
Db 241 ACTGACACTTGGCGCTCCCGCTCCAGGGCTCACCAGGGCTCTTTCAGACCTGGAGC 300
QY 301 TGCTTCTATTCTGGGACGTGAACCGCGAGCGGACTCCGCTGCTGAGCCTGTGC 360
```

```
Db 301 TGCTTCTATTCTGGGCGGTGAACCGCGAGCCGAGCTCGTTCCGCTGCTGTGAGCCTGTGC 360
QY 361 GGGGGGCTCCGCGAGGCTTTTGGCTACCGAGGCGCGAGTATGTCTATTAGCCCGCTGCC 420
Db 361 GGGGGGCTCCGCGAGGCTTTTGGCTACCGAGGCGCGAGTATGTCTATTAGCCCGCTGCC 420
QY 421 AATGTAGCGCGCGCGCGCGAGCGCAACAGCAGCGGCGCACACCTTCTCCAGCCCGG 480
Db 421 AATGTAGCGCGCGCGCGCGAGCGCAACAGCAGCGGCGCACACCTTCTCCAGCCCGG 480
QY 481 GGTGTTCGGGCGCGGCTTCCGGAGACCCACCTCTCGTTCGGGGTGGCTTCGGGGAG 540
Db 481 GGTGTTCGGGCGCGGCTTCCGGAGACCCACCTCTCGTTCGGGGTGGCTTCGGGGAG 540
QY 541 AACCCCGCATCTACGGGCGCTGGACCTTACAAAGCCGCGGCGGCGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTACGGGCGCTGGACCTTACAAAGCCGCGGCGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGAGGCTTGGGCGCGCAAGCGTTTCGTGTCTATCCCGGCTACGTG 660
Db 601 AGTCGTAGCGCGCGAGGCTTGGGCGCGCAAGCGTTTCGTGTCTATCCCGGCTACGTG 660
QY 661 GAGACCTGTGTTCGGGACGAGTCAATGGTCAAGTTCCACGGCGCGGACCTTGGAAAT 720
Db 661 GAGACCTGTGTTCGGGACGAGTCAATGGTCAAGTTCCACGGCGCGGACCTTGGAAAT 720
QY 721 TATCTCTGACGCTGCTGCAACGGCGCGGAGCTTACCGGCATCCAGCATCTCAAC 780
Db 721 TATCTCTGACGCTGCTGCAACGGCGCGGAGCTTACCGGCATCCAGCATCTCAAC 780
QY 781 CCCATCAACATCGTTGTGTCAAGGTGCTTCTTAGAGATCGTACCTCCCGGCGCAAG 840
Db 781 CCCATCAACATCGTTGTGTCAAGGTGCTTCTTAGAGATCGTACCTCCCGGCGCAAG 840
QY 841 GTACCGGCAATGCGGCGCTGACCGTTCGCTTCTTAGAGATCGTACCTCCCGGCGCAAG 900
Db 841 GTACCGGCAATGCGGCGCTGACCGTTCGCTTCTTAGAGATCGTACCTCCCGGCGCAAG 900
QY 901 AAAGTGTGACAGCACCCGAGTGTGGGACACTGCGGACCTGCTTCCACAGGCGAGAC 960
Db 901 AAAGTGTGACAGCACCCGAGTGTGGGACACTGCGGACCTGCTTCCACAGGCGAGAC 960
QY 961 CTGTGTGGAGCCACCTGTGACACCTTGGGCGGCTGATGGGTGATGATGATGATGATG 1020
Db 961 CTGTGTGGAGCCACCTGTGACACCTTGGGCGGCTGATGGGTGATGATGATGATGATG 1020
QY 1021 CCCAAGAGAGCTGCTGTGTCATTGAGGAGGATGGGCTTCCATCAGCCTTCAACCCTGCC 1080
Db 1021 CCCAAGAGAGCTGCTGTGTCATTGAGGAGGATGGGCTTCCATCAGCCTTCAACCCTGCC 1080
QY 1081 CACGAGCTGGGCGAGTGTTCACCATGCCCATGACATGTGAAGTGTGAAGTGTGTGAGAG 1140
Db 1081 CACGAGCTGGGCGAGTGTTCACCATGCCCATGACATGTGAAGTGTGAAGTGTGTGAGAG 1140
QY 1141 TTTGGGAAGTCCGAGCAACACATGATGTCCCCGAGCCCTCATCCAGATCGACCTGCC 1200
Db 1141 TTTGGGAAGTCCGAGCAACACATGATGTCCCCGAGCCCTCATCCAGATCGACCTGCC 1200
QY 1201 AACCCCTGTGAGCTGTGCTGCCATCATCCGACTTCTTGGAGAGGGG*ACGGT 1260
Db 1201 AACCCCTGTGAGCTGTGCTGCCATCATCCGACTTCTTGGAGAGGGG*ACGGT 1260
QY 1261 GACTGCTTCTTGGAGCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTCCCGGCGCC 1320
Db 1261 GACTGCTTCTTGGAGCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTCCCGGCGCC 1320
QY 1321 AGCTACACCTTGGAGCAGTGTGGCTTTTGGCGTGGGCTTCAAGCCCTGTCT 1380
Db 1321 AGCTACACCTTGGAGCAGTGTGGCTTTTGGCGTGGGCTTCAAGCCCTGTCT 1380
QY 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGAGGCAAGGACAGATGCTGTGC 1440
Db 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGAGGCAAGGACAGATGCTGTGC 1440
```

```
Db 1381 TACATGAGTACTGACCAAGCTGTGGTGACCGGGAAGGCAAGACAGATGGTGTGC 1440
Qy 1441 CAGACCGGCACATTCCTCGGCGCGATGGACACAGCTGTGGGAGGGCAAGCTCTGCCTC 1500
Db 1441 CAGACCGGCACATTCCTCGGCGCGATGGACACAGCTGTGGGAGGGCAAGCTCTGCCTC 1500
Qy 1501 AAAGGGGCTTGGCTGGAGAGACACAACCTCAACAAGACAG-----1541
Db 1501 AAAGGGGCTTGGCTGGAGAGACACAACCTCAACAAGACAGGAGGCAAT 1560
Qy 1542 -----GGTGGATGGT 1551
Db 1561 TCTCCAAAACAACCTTTATTAGGCTACCAATGGCTGCACACTACACAGGTGGATGGT 1620
Qy 1552 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGGCACATGTGGTGGGGCGTGCAG 1611
Db 1621 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGGCACATGTGGTGGGGCGTGCAG 1680
Qy 1612 CTGGCCAGGAGGAGTGCACCAACCCACCCCTGCCAACGGGGGCAAGTACTGCGAGGA 1671
Db 1681 CTGGCCAGGAGGAGTGCACCAACCCACCCCTGCCAACGGGGGCAAGTACTGCGAGGA 1740
Qy 1672 GTGAGGTGAATATCCGATCCTGGAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAG 1731
Db 1741 GTGAGGTGAATATCCGATCCTGGAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAG 1800
Qy 1732 AGCTTCGGGAGGAGCAGTGTAGGCTTTCAACGGCTTACACACAGCACAACCCGGCTC 1791
Db 1801 AGCTTCGGGAGGAGCAGTGTAGGCTTTCAACGGCTTACACACAGCACAACCCGGCTC 1860
Qy 1792 ACTCTCGCGGTGGATGGTGCCCAAGTACTCCGGCGGTGTCTCCGGGGACAAAGTCAAG 1851
Db 1861 ACTCTCGCGGTGGATGGTGCCCAAGTACTCCGGCGGTGTCTCCGGGGACAAAGTCAAG 1920
Qy 1852 CTCATCTGCCAGGCAATGGCACTGGCTACTTCTATGTCTGGCACCACCAAGTGGTGAC 1911
Db 1921 CTCATCTGCCAGGCAATGGCACTGGCTACTTCTATGTCTGGCACCACCAAGTGGTGAC 1980
Qy 1912 GGCAGCTGTCTCTCTGACTCCACCTCCGCTGTGTGTCTCAAGCAAGTGCATCAAGCT 1971
Db 1981 GGCAGCTGTCTCTCTGACTCCACCTCCGCTGTGTGTCTCAAGCAAGTGCATCAAGCT 2040
Qy 1972 GGCTGTGATGGGAACCTCGGCTCCAAGAAGATTCGACAAGTGTGGGGTGTGTGGGGGA 2031
Db 2041 GGCTGTGATGGGAACCTCGGCTCCAAGAAGATTCGACAAGTGTGGGGTGTGTGGGGGA 2100
Qy 2032 GACATAGAGCTCAAGAAGTGTGACTGTCTTCAACCAAGCCCATGCAATGCTACAAT 2091
Db 2101 GACATAGAGCTCAAGAAGTGTGACTGTCTTCAACCAAGCCCATGCAATGCTACAAT 2160
Qy 2092 TTCGTGGTGGCCATCCCGCAGGCGCCTCAAGCATCGACATCCGCCAGCGGTTACAAA 2151
Db 2161 TTCGTGGTGGCCATCCCGCAGGCGCCTCAAGCATCGACATCCGCCAGCGGTTACAAA 2220
Qy 2152 GGGCTGTATCGGGGATGACAATACCTGGCTCTGAAGAACGCCAAGCAAGTACCTGCTC 2211
Db 2221 GGGCTGTATCGGGGATGACAATACCTGGCTCTGAAGAACGCCAAGCAAGTACCTGCTC 2280
Qy 2212 AACGGCAATTTCTGTCTCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTG 2271
Db 2281 AACGGCAATTTCTGTCTCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTG 2340
Qy 2272 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAAGGCTTCCCGGCCCATCCTGGAG 2331
Db 2341 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAAGGCTTCCCGGCCCATCCTGGAG 2400
Qy 2332 CCGCTGACCGTGGAGGTCTCTCCGTGGGGAGATGACACCGCCCGGTTCCG 2385
Db 2401 CCGCTGACCGTGGAGGTCTCTCCGTGGGGAGATGACACCGCCCGGACCTCG 2454
```

RESULT 9
US-60-297-863-3

```
; Sequence 3, Application US/60297863
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses
; FILE OF INVENTION: Therefore
; CURRENT APPLICATION NUMBER: US/60/297,863
; CURRENT FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(2469)
US-60-297-863-3
```

```
Query Match 80.5%; Score 2298; DB 73; Length 2469;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

Qy 1 ATGCTTCTGCTGGGATCCTAACCTGGCTTTCCCGGGCGAAGCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCCTAACCTGGCTTTCCCGGGCGAAGCGCTGGAGGCTCTGAG 60
Qy 61 CCAGAGCGGGAGTGTAGTGTCCCATCCGACTGGACCCGAGCATTTAAGCGCGCGCTAC 120
Db 61 CCAGAGCGGGAGTGTAGTGTCCCATCCGACTGGACCCGAGCATTTAAGCGCGCGCTAC 120
Qy 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGCTCATTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGCTCATTTTCAGATCACAGCATTT 180
Qy 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGTCTAGTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGTCTAGTCTTGGCTCCCGCTTCTCC 240
Qy 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTACCGGGGCTCTTCAGAGCTGGAGCGC 300
Db 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTACCGGGGCTCTTCAGAGCTGGAGCGC 300
Qy 301 TGTCTTATCTGGGAGCTGAACCGCGAGCGGACTGTTGCTGCTGTGTAGGCTGTGC 360
Db 301 TGTCTTATCTGGGAGCTGAACCGCGAGCGGACTGTTGCTGCTGTGTAGGCTGTGC 360
Qy 361 GGGGGCTCCGCGGAGCTTTGGCTTACCGAGCGCGGATGTATGTCTATTAGCCCTGCC 420
Db 361 GGGGGCTCCGCGGAGCTTTGGCTTACCGAGCGCGGATGTATGTCTATTAGCCCTGCC 420
Qy 421 AATGCTAGCGCGCGCGGCGGAGCGCAACAGCGAGGGCGACACCTTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGCGGCGGAGCGCAACAGCGAGGGCGACACCTTCTCCAGCGCGG 480
Qy 481 GGTGTTCCGGGCGGGCTTCCGAGAGACCCACCTCTCGCTCGGGGTGGCTTGGGCTGG 540
Db 481 GGTGTTCCGGGCGGGCTTCCGAGAGACCCACCTCTCGCTCGGGGTGGCTTGGGCTGG 540
Qy 541 AACCCCGCCATCCTACGGGCTTGGAGCCCTTACAAGCCGCGCGGCGGGCTTCGGGAG 600
Db 541 AACCCCGCCATCCTACGGGCTTGGAGCCCTTACAAGCCGCGCGGCGGGCTTCGGGAG 600
Qy 601 AGTCGTAGCGCGCGAGTGTGGCGCGCAAGCTTGTGTCTATCCCGGCTACGCTG 660
Db 601 AGTCGTAGCGCGCGAGTGTGGCGCGCAAGCTTGTGTCTATCCCGGCTACGCTG 660
Qy 661 GAGACGCTGTGTGTGCGGACGAGTCAATGGTCAAGTTCACGCGCGGAGCAAT 720
Db 661 GAGACGCTGTGTGTGCGGACGAGTCAATGGTCAAGTTCACGCGCGGAGCAAT 720
Qy 721 TATCTGTGAGCTGTGGCAAGCGGGCGGAGCTACCGGCATCCGAGCATCTCTCAAC 780
Db 721 TATCTGTGAGCTGTGGCAAGCGGGCGGAGCTACCGGCATCCGAGCATCTCTCAAC 780
```

```
Db 721 TATCTGCTGACGCTGCTGCAACGGCGGAGCTCTACGGCCATCCCGCATCTCTCAAC 780
QY 781 CCATCAACATCGTTGCTGCTCAAGGTGCTGCTTCTAGAGATCTGAGTCCGGGCCAAG 840
Db 781 CCATCAACATCGTTGCTGCTCAAGGTGCTGCTTCTAGAGATCTGAGTCCGGGCCAAG 840
QY 841 GTACCGGCAATCGCGCCCTGACGGTGGCAACTTCTGTGCTTGGCAAGCAAGCTGAAC 900
Db 841 GTACCGGCAATCGCGCCCTGACGGTGGCAACTTCTGTGCTTGGCAAGCAAGCTGAAC 900
QY 901 AAGTGAAGTGAACGACCCGAGTACTGAGACACTGCCATCCGCTTCCACAGCAGGAC 960
Db 901 AAGTGAAGTGAACGACCCGAGTACTGAGACACTGCCATCCGCTTCCACAGCAGGAC 960
QY 961 CTGTGTGGAGCCACCATCTGTGACACCTGGGCATGGTGTGATG TGGTGTGATGTGAC 1020
Db 961 CTGTGTGGAGCCACCATCTGTGACACCTGGGCATGGTGTGATG TGGTGTGATGTGAC 1020
QY 1021 CCCAAGAGAGTGTCTGTCTGATTCAGGACGATGGCTTCCATAGGCTTCACCACTGCC 1080
Db 1021 CCCAAGAGAGTGTCTGTCTGATTCAGGACGATGGCTTCCATAGGCTTCACCACTGCC 1080
QY 1081 CAGGAGCTGGGCGAGCTGTTCAACATGCCCATGACATGTSAAGTCTGTGAGAGGTG 1140
Db 1081 CAGGAGCTGGGCGAGCTGTTCAACATGCCCATGACATGTSAAGTCTGTGAGAGGTG 1140
QY 1141 TTTGGGAAGCTCGGAGCCACACATGATGTCCCGGACCTCA TCCAGATCGACCGTGC 1200
Db 1141 TTTGGGAAGCTCGGAGCCACACATGATGTCCCGGACCTCA TCCAGATCGACCGTGC 1200
QY 1201 AACCCCTGTGACGCTGCACTGCTGCCATCATCACGACTTCC TGGCAGCGGCGACGGT 1260
Db 1201 AACCCCTGTGACGCTGCACTGCTGCCATCATCACGACTTCC TGGCAGCGGCGACGGT 1260
QY 1261 GACTGCTCTGAGCAACCCAGCAAGCCATCTCCCTGCCGAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTGAGCAACCCAGCAAGCCATCTCCCTGCCGAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCCCTGAGCAGCAGTGGAGCTGGCTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
Db 1321 AGCTACACCCCTGAGCAGCAGTGGAGCTGGCTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
QY 1381 TACATCAGTACTGACCAAGCTGTGTGACCCGGAGAGCCAGGACACAGATGGTGTGC 1440
Db 1381 TACATCAGTACTGACCAAGCTGTGTGACCCGGAGAGCCAGGACACAGATGGTGTGC 1440
QY 1441 CAGACCGGCACCTTCCCTGGCCGATGSCACAGCTGTGCGAGGGCAAGCTCTGCCTC 1500
Db 1441 CAGACCGGCACCTTCCCTGGCCGATGSCACAGCTGTGCGAGGGCAAGCTCTGCCTC 1500
QY 1501 AAGGGGCTGCGTGGAGAGACACAACTCAACAGCACAG- 1541
Db 1501 AAGGGGCTGCGTGGAGAGACACAACTCAACAGCACAG- 1560
QY 1542 -----GCTGATGTT 1551
Db 1542 -----GCTGATGTT 1551
QY 1561 TCTCCAAACAACTTTATAGGCTTACCAATGGCTTGCACATACACAGGTGGATGGT 1620
Db 1561 TCTCCAAACAACTTTATAGGCTTACCAATGGCTTGCACATACACAGGTGGATGGT 1620
QY 1552 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCACATGTGTGGGGGGTGCAG 1611
Db 1552 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCACATGTGTGGGGGGTGCAG 1611
QY 1621 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCACATGTGTGGGGGGTGCAG 1680
Db 1621 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCACATGTGTGGGGGGTGCAG 1680
QY 1612 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGACAGCGGGCAAGTACTGGAGGA 1671
Db 1612 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGACAGCGGGCAAGTACTGGAGGA 1671
QY 1681 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGACAGCGGGCAAGTACTGGAGGA 1740
Db 1681 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGACAGCGGGCAAGTACTGGAGGA 1740
QY 1672 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAG 1731
Db 1672 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAG 1731
QY 1741 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAG 1800
Db 1741 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAG 1800
QY 1732 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCACCAACCGGCTC 1791
Db 1732 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCACCAACCGGCTC 1791
QY 1801 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCACCAACCGGCTC 1860
Db 1801 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCACCAACCGGCTC 1860
```

```
QY 1792 ACTCTCGCGCTGGCATGGTGCCCAAGTACTCCGGCGCTGTCTCCCGGGACAAGTCAAG 1851
Db 1861 ACTCTCGCGCTGGCATGGTGCCCAAGTACTCCGGCGCTGTCTCCCGGGACAAGTCAAG 1920
QY 1852 CTCATCTCCGAGCAATGGCACTGGCTACTTCTATGTGTGGCACCACCAAGTGGTGGAC 1911
Db 1921 CTCATCTCCGAGCAATGGCACTGGCTACTTCTATGTGTGGCACCACCAAGTGGTGGAC 1980
QY 1912 GSCAGCTGTCTCTCTGACTCCACCTCGCTGTGTCTCAAGGCAAGTGCATCAAGGCT 1971
Db 1981 GSCAGCTGTCTCTCTGACTCCACCTCGCTGTGTCTCAAGGCAAGTGCATCAAGGCT 2040
QY 1972 GGCCTGTGATGGAACTCCCGCAGGCGCTCCAAAGAGATTTCGAAGTGTGGGGTGTGGGGGA 2031
Db 2041 GGCCTGTGATGGAACTCCCGCAGGCGCTCCAAAGAGATTTCGAAGTGTGGGGTGTGGGGGA 2100
QY 2032 GACATAGAGCTGCAAGAAGTGAAGTGTCTTACCAAGCCATGCATGCTACAAT 2091
Db 2101 GACATAGAGCTGCAAGAAGTGAAGTGTCTTACCAAGCCATGCATGCTACAAT 2160
QY 2092 TTCGTGTGGCCATCCCGCAGGCGCTCAAGCATCGACATTCGCCAGCGCGGTACAAA 2151
Db 2161 TTCGTGTGGCCATCCCGCAGGCGCTCAAGCATCGACATTCGCCAGCGCGGTACAAA 2220
QY 2152 GGCCTGTGATCGGGATGCAACTACTGCTCTGAAGAACAGCAAGTACCTGCTC 2211
Db 2221 GGCCTGTGATCGGGATGCAACTACTGCTCTGAAGAACAGCAAGTACCTGCTC 2280
QY 2212 AACGGGCAATTCGTGTGTGCTGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTG 2271
Db 2281 AACGGGCAATTCGTGTGTGCTGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTG 2340
QY 2272 CGGTACGCGCACGGGCGACAGCGCTGGAGAGCTTCAGGCTTCCCGGCCATCTCGAG 2331
Db 2341 CGGTACGCGCACGGGCGACAGCGCTGGAGAGCTTCAGGCTTCCCGGCCATCTCGAG 2400
QY 2332 CGGCTGACCGTGGAGGTCTCTCCGTGGGGAAGTGAACCGCCCGCGGGTCCCG 2385
Db 2401 CGGCTGACCGTGGAGGTCTCTCCGTGGGGAAGTGAACCGCCCGCGGGTCCCG 2454
```

```
RESULT 10
US-10-163-316-1
; Sequence 1, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI01-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940)
; OTHER INFORMATION: n = A,T,C or G
US-10-163-316-1
```

```
Query Match 80.5%; Score 2298; DB 41; Length 2940;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Caps 1;
```


QY 1501 AAAGGGCCCTGCTGGAGAGACAAACCTCAACAAGCACAG----- 1541
Db 1972 AAAGGGCCCTGCTGGAGAGACAAACCTCAACAAGCACAGGCTCGACTGACATCAT 2031
QY 1542 -----GGTGATGCT 1551
Db 2032 TCTCAAAACAACCTTTATTAAAGGCTACCAAAATGGCTGCACACTACACAGTGGATGCT 2091
QY 1552 TCCTGGCCCAATGGATCCCTATGCGCCCTGCTGCGGCACATGCTGGGGCGCTGCAG 1611
Db 2092 TCCTGGCCCAATGGATCCCTATGCGCCCTGCTGCGGCACATGCTGGGGCGCTGCAG 2151
QY 1612 CTGGCCAGGAGGAGTGCACCAACCCACCCCTGCGCAACGGGGCAAGTACTCGAGGGA 1671
Db 2152 CTGGCCAGGAGGAGTGCACCAACCCACCCCTGCGCAACGGGGCAAGTACTCGAGGGA 2211
QY 1672 GTGAGGTGAATACCGATCCTGCAATCTGGAGGCTGCGCCAGCTAGCTCGCGAAAG 1731
Db 2212 GTGAGGTGAATACCGATCCTGCAATCTGGAGGCTGCGCCAGCTAGCTCGCGAAAG 2271
QY 1732 AGCTTCGGGAGGAGGAGTGTGAGGCTTTCAACGGCTTACACACAGCACCAACCGGCTC 1791
Db 2272 AGCTTCGGGAGGAGGAGTGTGAGGCTTTCAACGGCTTACACACAGCACCAACCGGCTC 2331
QY 1792 ACTCTCGCCGTGGGATGGGTGCCAAAGTACTCGGGGCTGTCTCCCGGGACAAAGTCCAAG 1851
Db 2332 ACTCTCGCCGTGGGATGGGTGCCAAAGTACTCGGGGCTGTCTCCCGGGACAAAGTCCAAG 2391
QY 1852 CTCATCTCGCGAGCAATGACATGGCTACTTCTATGCTGTCGACCAAGTGGTGGAC 1911
Db 2392 CTCATCTCGCGAGCAATGACATGGCTACTTCTATGCTGTCGACCAAGTGGTGGAC 2451
QY 1912 GGCACGTGTGCTCTCTGACTCCACTCGTCTGTGTCTCAAGGCAAGTGCATCAAGGCT 1971
Db 2452 GGCACGTGTGCTCTCTGACTCCACTCGTCTGTGTCTCAAGGCAAGTGCATCAAGGCT 2511
QY 1972 GGCTGTGATGGGAACCTGGGCTCCAGAAGAGATTCGACAAAGTGTGGGGTGTGTGGGGA 2031
Db 2512 GGCTGTGATGGGAACCTGGGCTCCAGAAGAGATTCGACAAAGTGTGGGGTGTGTGGGGA 2571
QY 2032 GACAAATAGAGCTCAAGAGAGTGTGACTGCTCTTCAACAGCCCATGCTGCTACAAAT 2091
Db 2572 GACAAATAGAGCTCAAGAGAGTGTGACTGCTCTTCAACAGCCCATGCTGCTACAAAT 2631
QY 2092 TTCGTGTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGGCGGTTACAAA 2151
Db 2632 TTCGTGTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGGCGGTTACAAA 2691
QY 2152 GGGCTGATCGGGATGACAACTACCTGGCTCTGAGAACAGCAAGCAAGTACCTGCTC 2211
Db 2692 GGGCTGATCGGGATGACAACTACCTGGCTCTGAGAACAGCAAGCAAGTACCTGCTC 2751
QY 2212 AACGGCATTTCTGTGTGCTGGCGGTGGAGCGGACCTGTGTGGTGAAGGCAAGTGTGCTG 2271
Db 2752 AACGGCATTTCTGTGTGCTGGCGGTGGAGCGGACCTGTGTGGTGAAGGCAAGTGTGCTG 2811
QY 2272 CGGTACAGCGCACGGGACAGCGGTGGAGGCTGTGAGGCTTCCCGGCGCCATCTCTGGAG 2331
Db 2812 CGGTACAGCGCACGGGACAGCGGTGGAGGCTGTGAGGCTTCCCGGCGCCATCTCTGGAG 2871
QY 2332 CCGTGTACCGGTGAGGCTCTCTCGTGGGGAAGATGACACCGCTCCGGGTCCCG 2385
Db 2872 CCGTGTACCGGTGAGGCTCTCTCGTGGGGAAGATGACACCGCTCCGGGTCCCG 2925

RESULT 12
US-60-212-656-762
; Sequence 762, Application US/60212656
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000673
; CURRENT APPLICATION NUMBER: US/60/212,656
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 762
; LENGTH: 3400
; TYPE: DNA
; ORGANISM: HUMAN
; US-60-212-656-762

Query Match 80.1%; Score 2285.4; DB 65; Length 3400;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 545; Gaps 2;

QY 1 ATGCTTCTGCTGGGATCCTAAACCTGCTTTCGCCGGGCAACCGCTGGAGGCTCTGAG 60
Db 2 ATGCTTCTGCTGGGATCCTAAACCTGCTTTCGCCGGGCAACCGCTGGAGGCTCTGAG 61
QY 61 CCAGAGCGGGAGTAGTCTGTTCCCATCGACTGGACCCGGACATTAAACGGCCGCCGCTAC 120
Db 62 CCAGAGCGGGAGTAGTCTGTTCCCATCGACTGGACCCGGACATTAAACGGCCGCCGCTAC 121
QY 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT 180
Db 122 TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT 181
QY 181 CAGGAGGACTTTTACTACTACGCTGACGCGGATGCTCAAGTTCTTGGTCCCGCTCTCC 240
Db 182 CAGGAGGACTTTTACTACTACGCTGACGCGGATGCTCAAGTTCTTGGTCCCGCTCTCC 241
QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTACAGACTGCGAGGC 300
Db 242 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTACAGACTGCGAGGC 301
QY 301 TCTTCTTATCTGGGAGCTGAACCCGAGCGGACTGTTCTGCTGCTGTGAGC-TGTGC 360
Db 302 TCTTCTTATCTGGGAGCTGAACCCGAGCGGACTGTTCTGCTGCTGTGAGC-TGTGC 361
QY 361 GGGGGCTCCCGGAGCCCTTTGGCTTACCGAGGCGCGGAGTAGTATGTAATAGCCCG-TGCC 420
Db 362 GGGGGCTCCCGGAGCCCTTTGGCTTACCGAGGCGCGGAGTAGTATGTAATAGCCCG-TGCC 421
QY 421 AATGCTAGCGCGCGCGGCGGAGCAACAGCGAGCGGACACCTTCTCCAGCGCGG 480
Db 422 AATGCTAGCGCGCGCGGCGGAGCAACAGCGAGCGGACACCTTCTCCAGCGCGG 481
QY 481 GGTGTTCCGGGCGGCTTCCGGAGACCCACCTCTCTGCTGCGGGGTGGCTCGGGCTGG 540
Db 482 GGTGTTCCGGGCGGCTTCCGGAGACCCACCTCTCTGCTGCGGGGTGGCTCGGGCTGG 541
QY 541 AACCCCGCCATCCTACGGGCGCTTGACCCCTTACAAAGCGCGGCGGCGGCTTCGGGAG 600
Db 542 AACCCCGCCATCCTACGGGCGCTTGACCCCTTACAAAGCGCGGCGGCGGCTTCGGGAG 601
QY 601 AGTCTAGCGCGCGGAGTCTGGGCGCGGAGCGCTTTCGCTATCTCCCGGTAACGTG 660
Db 602 AGTCTAGCGCGCGGAGTCTGGGCGCGGAGCGCTTTCGCTATCTCCCGGTAACGTG 661
QY 661 GAGAGCTGCTGCTCGGAGCAGTCAATGTGTAAGTTCCAGCGCGGAGCTTGAACAT 720
Db 662 GAGAGCTGCTGCTCGGAGCAGTCAATGTGTAAGTTCCAGCGCGGAGCTTGAACAT 721
QY 721 TATCTGCTGACGCTGCTGCGAACGCGGCGGAGTCTACCGGCATCCAGAGTCTCAAC 780
Db 722 TATCTGCTGACGCTGCTGCGAACGCGGCGGAGTCTACCGGCATCCAGAGTCTCAAC 781
QY 781 CCCATCAACATGCTGCTGTAAGTGTGCTTCTTAGAGATCTGACTCCGGGCGGCAAG 840
Db 782 CCCATCAACATGCTGCTGTAAGTGTGCTTCTTAGAGATCTGACTCCGGGCGGCAAG 841
QY 841 GTCACCGGCAATCGCGGCCCTGACGCTGCGCAACTTCTGTGCTGCGCAAGAAGTGAAC 900

Db 842 GTACCGGCAATGCGCCCTGACGTGCGCAACTTCTGTGCTGCGAAGAAGCTGAAC 901
QY 901 AAAGTCAGTGACAAGCACCCGAGTACTGGGACACTGCCATCCTCTTACC----- 951
Db 902 AAAGTCAGTGACAAGCACCCGAGTACTGGGACACTGCCATCCTCTTACCAGGCGAGTT 961
QY 952 ----- 951
Db 962 CCAGGGGAGAGCCCTCTCCACGCTCCGCGGAGCGGCGCTACAGTGCATCTGGGCCATT 1021
QY 952 ----- 951
Db 1022 GGAGGAGAGCCCTGCGCTTCCGAAGGTGTGGCCCTGGGCGGCCAATCAGCGCCCTCTGG 1081
QY 952 ----- 951
Db 1082 ATCAGGCGCGAGGGCGGAACCCAGGAAGTTGCCGCGCCTGGAGCTGCACTTGTGTGCC 1141
QY 952 ----- 951
Db 1142 AAGACCGATAGGAGACGCGGTGAGGATGTGTGGAGAGCGGGAAGCGGCCACCCCTA 1201
QY 952 ----- 951
Db 1202 TTGTATGGCGGCTGAGTCTTCTCGGACACCTCCTGAGGTCTGCTTTCAAGGTGTAGA 1261
QY 952 ----- 951
Db 1262 ACTGAAGGTGATCCAAAGTGTGGCTCCCAATCCTAAATTTATTTAAGAAAGGACAGTTA 1321
QY 952 ----- 951
Db 1322 GAAATAACAGTCCAGGAGATTAAAGTTCTGTGGCCAGAGATGATTCAGGATCATCTGGA 1381
QY 952 ----- 951
Db 1382 AAAGTGGCTGCAGCTTACAAAGAAAACAGCCAAATGTGCGAATCGTGAGAGCAGGACCTG 1441
QY 964 TGTGGAGCCACACCTGTGACACCTGGCATGTGCTGATGTGGGTACCATCTGTGACCCC 1023
Db 1442 TGTGGAGCCACACCTGTGACACCTGGCATGTGCTGATGTGGGTACCATCTGTGACCCC 1501
QY 1024 AAGAAAGCTGTCTGTGCTATTGAGAGAGATGGGCTTCCATCAGCCTTCACTGCGCAC 1083
Db 1502 AAGAAAGCTGTCTGTGCTATTGAGAGAGATGGGCTTCCATCAGCCTTCACTGCGCAC 1561
QY 1084 GAGCTGGGCCACGTGTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTGTT 1143
Db 1562 GAGCTGGGCCACGTGTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTGTT 1621
QY 1144 GGAAGCTCCGAGCCAAACCATGATGTCCCGGACCTCATCCAGATCGACCGTGGCAAC 1203
Db 1622 GGAAGCTCCGAGCCAAACCATGATGTCCCGGACCTCATCCAGATCGACCGTGGCAAC 1681
QY 1204 CCCTGGTCAGCTGAGTGTGCCATCATCACCAGCTTCTTGGACAGCGGCGAGGTGAC 1263
Db 1682 CCCTGGTCAGCTGAGTGTGCCATCATCACCAGCTTCTTGGACAGCGGCGAGGTGAC 1741
QY 1264 TGCCTCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCAGGATCTGCCGGGCGGCGAGC 1323
Db 1742 TGCCTCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCAGGATCTGCCGGGCGGCGAGC 1801
QY 1324 TACACCTTGAGCCAGCAGTGGAGCTTGTGGCTGGGCTCCAAAGCCCTGTGCTTAC 1383
Db 1802 TACACCTTGAGCCAGCAGTGGAGCTTGTGGCTGGGCTCCAAAGCCCTGTGCTTAC 1861
QY 1384 ATGAGTACTGCACCAAGCTGTGTGACCGGGAGGCCAAGGACAGATGGTGTGCCAG 1443
Db 1862 ATGAGTACTGCACCAAGCTGTGTGACCGGGAGGCCAAGGACAGATGGTGTGCCAG 1921
QY 1444 ACCCGCACCTTCCCTGGCGGATGACACAGCTGTGGCGAGGCGCAAGCTCTGCCTCAA 1503

Db 1922 ACCCGCACCTTCCCTGGCGGATGGCACCAAGCTGTGGCGAGGCGCAAGCTCTGCCTCAA 1981
QY 1504 GGGGCTTGGTGGAGAGACAACTCAACAAGCACAG----- 1541
Db 1982 GGGGCTTGGTGGAGAGACAACTCAACAAGCACAGGGCTTTGACTGACATCATTTCT 2041
QY 1542 -----GGTGGATGGTTC 1554
Db 2042 CCAAAACAACCTCTATTAAAGGCTACCAATGGGCTGCACACTAGAAAGTGGATGGTTC 2101
QY 1555 TGGCCAAATGGGATCCTATGGCCCTGTGCGGCGACATGTGTGGGGGCGGTGCAACTG 1614
Db 2102 TGGCCAAATGGGATCCTATGGCCCTGTGCGGCGACATGTGTGGGGGCGGTGCAACTG 2161
QY 1615 GCCAGGAGGAGTGCACCAACCCCTGCGCCACCGGGGGCAAGTACTGCGAGGGAGT 1674
Db 2162 GCCAGGAGGAGTGCACCAACCCCTGCGCCACCGGGGGCAAGTACTGCGAGGGAGT 2221
QY 1675 AGGTGAAATACCGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCTCCGGAAAGAGC 1734
Db 2222 AGGTGAAATACCGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCTCCGGAAAGAGC 2281
QY 1735 TTCCGGAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACACAGCACCAACCGGCTCAT 1794
Db 2282 TTCCGGAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACACAGCACCAACCGGCTCAT 2341
QY 1795 CTGCGCTGTCATGGGTGCGCAAGTACTCCGGGCTGTCTCCCGGGGACAAAGTGCACACTC 1854
Db 2342 CTGCGCTGTCATGGGTGCGCAAGTACTCCGGGCTGTCTCCCGGGGACAAAGTGCACACTC 2401
QY 1855 ATCTGCGGAGCCAAATGCGCTACTTCTATGTGCTGCGCACCCAGGTGGTGGACGC 1914
Db 2402 ATCTGCGGAGCCAAATGCGCTACTTCTATGTGCTGCGCACCCAGGTGGTGGACGC 2461
QY 1915 ACCTGTGCTCTCCTGACTCCACCTCGTCTGTGTCAGGCAAGTGCATCAAGGCTGGC 1974
Db 2462 ACCTGTGCTCTCCTGACTCCACCTCGTCTGTGTCAGGCAAGTGCATCAAGGCTGGC 2521
QY 1975 TGTGATGGGAACCTGGGCTCCAAAGAGAGATTCGCAAGTGTGGGTGTGTGGGGGAGAC 2034
Db 2522 TGTGATGGGAACCTGGGCTCCAAAGAGAGATTCGCAAGTGTGGGTGTGTGGGGGAGAC 2581
QY 2035 AATAAGAGCTGCAAGAGGTGACTGACTCTTCCAAAGCCCATGTCATGCTACAAATTC 2094
Db 2582 AATAAGAGCTGCAAGAGGTGACTGACTCTTCCAAAGCCCATGTCATGCTACAAATTC 2641
QY 2095 GTGGTGGCCATCCCGAGCGGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGG 2154
Db 2642 GTGGTGGCCATCCCGAGCGGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGG 2701
QY 2155 CTGATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCAAGTACCTGCTCAA 2214
Db 2702 CTGATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCAAGTACCTGCTCAA 2761
QY 2215 GGCATTTCTGTGTGCTGGCGGTGGAGCGGACCTGCTGTTGAAGGCGAGTCTGTGCGG 2274
Db 2762 GGCATTTCTGTGTGCTGGCGGTGGAGCGGACCTGCTGTTGAAGGCGAGTCTGTGCGG 2821
QY 2275 TACAGCGGACGGGACAGCGGTGGAGAGCTTCCGCGGCCATCTCTGAGAGCG 2334
Db 2822 TACAGCGGACGGGACAGCGGTGGAGAGCTTCCGCGGCCATCTCTGAGAGCG 2881
QY 2335 CTGACCTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGTCCGCTACTCTCTTC 2394
Db 2882 CTGACCTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGTCCGCTACTCTCTTC 2941
QY 2395 TATCTGCCCAAAGAGCTCGGGAGGACAAAGTCTCTCATCCCAAGACACCCCGGAGCC 2454
Db 2942 TATCTGCCCAAAGAGCTCGGGAGGACAAAGTCTCTCATCCCAAGACACCCCGGAGCC 3001
QY 2455 TCTGTCTTGCACACAGCGTCTCTCAGCTCTCCAAACAGGTGGAGAGCGGACACAGG 2514
Db 3002 TCTGTCTTGCACACAGCGTCTCTCAGCTCTCCAAACAGGTGGAGAGCGGACACAGG 3061

QY 2515 CCCCTGACGCTGGGTGGCTGAGCTGGGGCCGCTGCTCCGAGCTGGCGCAGTGGC 2574
Db 3062 CCCCTGACGCTGGGTGGCTGAGCTGGGGCCGCTGCTCCGAGCTGGCGCAGTGGC 3121
QY 2575 CTCAGAAAGCGGCGGTGGACTGCGGGGCTCGCGGGGACGCGACGCTGCGCTGT 2634
Db 3122 CTCAGAAAGCGGCGGTGGACTGCGGGGCTCGCGGGGACGCGACGCTGCGCTGT 3181
QY 2635 GATGCAGCCCATCGGCGCGTGGAGACACAAAGCTGCGGGGACGCGCTGCGCCACCTGGGAG 2694
Db 3182 GATGCAGCCCATCGGCGCGTGGAGACACAAAGCTGCGGGGACGCGCTGCGCCACCTGGGAG 3241
QY 2695 CTCAGCGCCTGGTCAACCTGCTCAAGAGCTGCGGGCGCGGATTCAGAGCGGCTCACATC 2754
Db 3242 CTCAGCGCCTGGTCAACCTGCTCAAGAGCTGCGGGCGCGGATTCAGAGCGGCTCACATC 3301
QY 2755 AAGTGTGTGGGCGACGAGGCGCGGCTGCTGGGCGCGGACAGTCAACTTGCACGCAAG 2814
Db 3302 AAGTGTGTGGGCGACGAGGCGCGGCTGCTGGGCGCGGACAGTCAACTTGCACGCAAG 3361
QY 2815 CCCAGAGCTGACTTCTCGTCTCCTCAGGCGCTGCTGA 2853
Db 3362 CCCAGAGCTGACTTCTCGTCTCCTCAGGCGCTGCTGA 3400

RESULT 13

US-60-242-679-1814
; Sequence 1814, Application US/60242679
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steven Istvan
; APPLICANT: Spier, Eugene
; APPLICANT: Greenberg, Simon
; APPLICANT: Brandenberger, Ralph
; APPLICANT: Wang, Yu
; APPLICANT: Dubman, Alex
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000898-PROV
; CURRENT APPLICATION NUMBER: US/60/242, 679
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 2265
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1814
; LENGTH: 3759
; TYPE: DNA
; ORGANISM: HUMAN
US-60-242-679-1814

Query Match 78.0%; Score 2225.4; DB 68; Length 3759;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 606; Gaps 2;
QY 1 ATGCTTCTGCTGGCATCCTAACCTGGCTTTCGCCGGGGAACCGCTGAGGCTCTGAG 60
Db 301 ATGCTTCTGCTGGCATCCTAACCTGGCTTTCGCCGGGGAACCGCTGAGGCTCTGAG 360
QY 61 CCAGAGGGGAGGTAGTGTTCCTCCACTGGACCCGGACATTAACGGCGCGCGCTAC 120
Db 361 CCAGAGGGGAGGTAGTGTTCCTCCACTGGACCCGGACATTAACGGCGCGCGCTAC 420
QY 121 TACTGGCGGGTCCCGAGGACTCGGGGATCAGGACTCATTTTCAGATCAGACATTT 180
Db 421 TACTGGCGGGTCCCGAGGACTCGGGGATCAGGACTCATTTTCAGATCAGACATTT 480
QY 181 CAGGAGGACTTTTACCTACACTGACCCGGATGCTCAGTCTTGGCTCCCGCTCTCC 240
Db 481 CAGGAGGACTTTTACCTACACTGACCCGGATGCTCAGTCTTGGCTCCCGCTCTCC 540
QY 241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGCCAGCG 300
Db 541 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGCCAGCG 600

QY 301 TGCCTTATTTCTGGGACGTGAACCCGAGCCGAGCTCGTTTCGCTGCTGTAGCCTGTGC 360
Db 601 TGCCTTATTTCTGGGACGTGAACCCGAGCCGAGCTCGTTTCGCTGCTGTAGCCTGTGC 660
QY 361 GGGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCCGAGTATGTCAATAGCCGCTGCC 420
Db 661 GGGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCCGAGTATGTCAATAGCCGCTGCC 720
QY 421 AATGCTAGCGCGCGCGGCGCAGCGAACAGCCAGCGCGCACACCTTCTCCAGCGCGG 480
Db 721 AATGCTAGCGCGCGCGGCGCAGCGAACAGCCAGCGCGCACACCTTCTCCAGCGCGG 780
QY 481 GGTGTTCGCGGCGGCGCTTCGCGAGACCCACCTCTCGCTGCGGGGTGGGCTCGGGCTGG 540
Db 781 GGTGTTCGCGGCGGCGCTTCGCGAGACCCACCTCTCGCTGCGGGGTGGGCTCGGGCTGG 840
QY 541 AACCOCGCCATCTTACGGGCGCTTGACGCCCTTACAAGCGCGCGGCGGCTTCGCGGAG 600
Db 841 AACCOCGCCATCTTACGGGCGCTTGACGCCCTTACAAGCGCGCGGCGGCTTCGCGGAG 900
QY 601 AGTCTAGCGCGCGCAGGCTCTGGCGCGGCCAAGCGTTTCTGTCTATCCCGCGTACGTG 660
Db 901 AGTCTAGCGCGCGCAGGCTCTGGCGCGGCCAAGCGTTTCTGTCTATCCCGCGTACGTG 960
QY 661 GAGACGCTGGTGTGGCGGAGAGTCAATGTCAAGTTCCACGCGCGGAGCTTGAACAT 720
Db 961 GAGACGCTGGTGTGGCGGAGAGTCAATGTCAAGTTCCACGCGCGGAGCTTGAACAT 1020
QY 721 TATCTGCTGAGCTGCTGCGCAACGCGCGGCTTACCGCCATCCAGCATCTCTCAAC 780
Db 1021 TATCTGCTGAGCTGCTGCGCAACGCGCGGCTTACCGCCATCCAGCATCTCTCAAC 1080
QY 781 CCATCAACATCGTTGTGTCAAGGTGCTCTTTAGAGATCGTGACTCGCGGCCCAAG 840
Db 1081 CCATCAACATCGTTGTGTCAAGGTGCTCTTTAGAGATCGTGACTCGCGGCCCAAG 1140
QY 841 GTACCGGCAATCGCGGCTGACGCTGCGCACTTCTGTGCTGCGAGCAAGCAAGTGAAC 900
Db 1141 GTACCGGCAATCGCGGCTGACGCTGCGCACTTCTGTGCTGCGAGCAAGCAAGTGAAC 1200
QY 901 AAAGTGAGTGACAACGACCCGAGTACTGGGACACTGCGCATCTTCCACGAGCA --- 956
Db 1201 AAAGTGAGTGACAACGACCCGAGTACTGGGACACTGCGCATCTTCCACGAGCA --- 1260
QY 957 --- --- --- --- --- 956
Db 1261 CCAGGGGAGAGCCCTCTCCACGCTCCCGGAGCGCGGCTCACGTGCATCTGGGCCATT 1320
QY 957 --- --- --- --- --- 956
Db 1321 GGAGGAGAGCCTGCGCTTTCGAAAGGTGTGGCCTGGCGGCGCCANTCAGCGCCTCTG 1380
QY 957 --- --- --- --- --- 956
Db 1381 ATCAGGCGCCGAGSGGCGGAAACCAGGAAAGTTGCCGCCCGGAGCTGCAAGTTTGTC 1440
QY 957 --- --- --- --- --- 956
Db 1441 AAGACGATAGGAGACGCGGTGAGGATGTTGGAGAGGCGGAGCGGCCACCCCTA 1500
QY 957 --- --- --- --- --- 956
Db 1501 TTGTATGGCGGCTGAGTCTTTCGAGACACTTCTCGGACACTTCTTCAAGGGTTGTAGA 1560
QY 957 --- --- --- --- --- 956
Db 1561 ACTGAAGGCTCTCCCGACTCAGTGTCAACCGGCACCTTGTGAGCTACAGGGCGAGGCC 1620
QY 957 --- --- --- --- --- 972
Db 1621 TTCCATTCTTTGGGGGATATGACACGAGGAGGACAGGAGTCAATGAGACTGTGTGAGGCC 1680

QY 973 ACCACCTGTGACACCCCTGGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAGC 1032
Db 1681 ACCACCTGTGACACCCCTGGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAGC 1740
QY 1033 TGCTCTGTGATTCAGGACGATGGGCTTCATCAGCCCTTACACACTGCCCCAAGAGTGGGC 1092
Db 1741 TGCTCTGTGATTCAGGACGATGGGCTTCATCAGCCCTTACACACTGCCCCAAGAGTGGGC 1800
QY 1093 CACGTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGAGGTGTTTGGGAAGCTC 1152
Db 1801 CACGTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGAGGTGTTTGGGAAGCTC 1860
QY 1153 CGAGCAACCCACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGCCACCCCTGTGTCA 1212
Db 1861 CGAGCAACCCACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGCCACCCCTGTGTCA 1920
QY 1213 GCCTGCAGTGTGCCATCATCACCAGCTTCCCTGGACAGCGGGGACGGTGAAGTGCCTCCTG 1272
Db 1921 GCCTGCAGTGTGCCATCATCACCAGCTTCCCTGGACAGCGGGGACGGTGAAGTGCCTCCTG 1980
QY 1273 GACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGCCAGCTACACCCCTG 1332
Db 1981 GACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGCCAGCTACACCCCTG 2040
QY 1333 AGCCAGAGTGGAGCTGGCTTTTGGCGTGGGCTCCAGCCCTGTCCCTTACATGCAATAC 1392
Db 2041 AGCCAGAGTGGAGCTGGCTTTTGGCGTGGGCTCCAGCCCTGTCCCTTACATGCAATAC 2100
QY 1393 TGCACCAAGCTGTGTGCACCGGAGGCGCAAGGACAGATGTTGTGCCAGACCCGCCAC 1452
Db 2101 TGCACCAAGCTGTGTGCACCGGAGGCGCAAGGACAGATGTTGTGCCAGACCCGCCAC 2160
QY 1453 TTCCCTCGGGCGGATGGCACCAGCTGTGGCGAGGGCAAGCTCTCCCTCAAAGGGGCGCTGC 1512
Db 2161 TTCCCTCGGGCGGATGGCACCAGCTGTGGCGAGGGCAAGCTCTCCCTCAAAGGGGCGCTGC 2220
QY 1513 GTGAGAGACACACCTCAACACACAG----- 1541
Db 2221 GTGAGAGACACACCTCAACACACAGGCTTTCACTGATGACCTTTCTCTCATTTCC 2280
QY 1542 ----- 1541
Db 2281 CGACGGGGGAAAAACCAACAGGTTCCCGGAGGAGTTGTACAGCTTGTGTCTCCAAT 2340
QY 1542 ----- 1541
Db 2341 TCCATCTTCCCATTTTCTCTGCTGCCCATTCACACAGGCTTTGACTGACATCAATTC 2400
QY 1542 -----GGTGGATGGTTCC 1554
Db 2401 CCAAAACACATCTATTAAAGCTACCAAAATGGCCTGCACACTACGAAGGTTGGATGGTTCC 2460
QY 1555 TGGCCCAATGGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGGGCGGTGCGAGCTG 1614
Db 2461 TGGCCCAATGGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGGGCGGTGCGAGCTG 2520
QY 1615 GCGAGGAGGAGTGCACACACCCACCCCTGCGCAAGGGGGCAAGTACTGCGAGGAGTG 1674
Db 2521 GCGAGGAGGAGTGCACACACCCACCCCTGCGCAAGGGGGCAAGTACTGCGAGGAGTG 2580
QY 1675 AGGTGAAATACCGATCTGCAATCTGAGCCCTGAGCCCTGCCAAGCGGGGCAAGTACTGCGAGGAGTG 1734
Db 2581 AGGTGAAATACCGATCTGCAATCTGAGCCCTGAGCCCTGCCAAGCGGGGCAAGTACTGCGAGGAGTG 2640
QY 1735 TTCGGGAGGAGGAGTGTGAGGCTTTTCAAGCGGTACAACACAGCACCACAGTGGCTCACT 1794
Db 2641 TTCGGGAGGAGGAGTGTGAGGCTTTTCAAGCGGTACAACACAGCACCACAGTGGCTCACT 2700
QY 1795 CTGCGCTGGGATGGGTGCCAAGTACTCGGGCTGTCTCCCGGGGACAAAGTGCAGAGCTC 1854
Db 2701 CTGCGCTGGGATGGGTGCCAAGTACTCGGGCTGTCTCCCGGGGACAAAGTGCAGAGCTC 2760
QY 1855 ATCTGCCGAGCCCAATGGCACTGGCTACTTCTATGTGCTGGCACCACCAAGGTTGGTGGACGGC 1914

Db 2761 ATCTGCCGAGCCCAATGGCACTGGCTACTTCTATGTGCTGGCACCACCAAGTGGTGGAGTGGC 2820
QY 1915 ACCTGTGCTCTCTCTGACTCCACTCCGCTCTGTGTCAAGGCAAGTGCATCAAGGCTGGC 1974
Db 2821 ACCTGTGCTCTCTCTGACTCCACTCCGCTCTGTGTCAAGGCAAGTGCATCAAGGCTGGC 2880
QY 1975 TGTGATGGGAACCTGGGCTCCAAAGAGAGATTGCAAGTGTGGGTGTGTGGGGGAGAC 2034
Db 2881 TGTGATGGGAACCTGGGCTCCAAAGAGAGATTGCAAGTGTGGGTGTGTGGGGGAGAC 2940
QY 2035 AATAAGAGCTGCAAGAGGTTGACTGGACTCTTCCACCAAGCCCATGATGGCTACAAATTC 2094
Db 2941 AATAAGAGCTGCAAGAGGTTGACTGGACTCTTCCACCAAGCCCATGATGGCTACAAATTC 3000
QY 2095 GTGGTGGCCATCCCGCAGCGCTCAAGCATCGACATCCCGCAGCGCGGTACAAAGGG 2154
Db 3001 GTGGTGGCCATCCCGCAGCGCTCAAGCATCGACATCCCGCAGCGCGGTACAAAGGG 3060
QY 2155 CTGATCGGGATGACAACTACCTTGGCTCTGGAAGACAGCCAAAGTACCTGTAAAC 2214
Db 3061 CTGATCGGGATGACAACTACCTTGGCTCTGGAAGACAGCCAAAGTACCTGTAAAC 3120
QY 2215 GGGCATTTCTGTGTGTGGGGTGGAGCGGACTCTGTGTGAAGGCGAGTCTGTCCGG 2274
Db 3121 GGGCATTTCTGTGTGTGGGGTGGAGCGGACTCTGTGTGAAGGCGAGTCTGTCCGG 3180
QY 2275 TACAGCGGCACGGGCACAGCGTGGAGAGCTTCCAGGCTTCCCGGCCCATCTCTGAAATTCG 2334
Db 3181 TACAGCGGCACGGGCACAGCGTGGAGAGCTTCCAGGCTTCCCGGCCCATCTCTGAAATTCG 3240
QY 2335 CTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTTC 2394
Db 3241 CTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTTC 3300
QY 2395 TATCTGCCAAAGAGCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCC 2454
Db 3301 TATCTGCCAAAGAGCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCC 3360
QY 2455 TCTGTCTTGCACAAACAGCTCTCAGCCCTTCCAAACAGGTGGAGAGCGGACACAGG 2514
Db 3361 TCTGTCTTGCACAAACAGCTCTCAGCCCTTCCAAACAGGTGGAGAGCGGACACAGG 3420
QY 2515 CCCCCCTGACGCTGGGTGGCTGGAGCTGGGGCGTGTCCGCGAGCTGCGGAGTGGCAGTGGC 2574
Db 3421 CCCCCCTGACGCTGGGTGGCTGGAGCTGGGGCGTGTCTCCGAGCTTGGCGGAGTGGCAGTGGC 3480
QY 2575 CTGCAAGAGCGGGGCTGGACTTCTCGGGCTCCCGGGGCGAGCGACGCTCCCTTGT 2634
Db 3481 CTGCAAGAGCGGGGCTGGACTTCCCGGGGCTCCCGGGGCGAGCGACGCTCCCTTGT 3540
QY 2635 GATGAGCCCATCGGCGCTGGAGACACAAAGCTTGGGGGAGCCCTGCCCCACCTTGTAG 2694
Db 3541 GATGAGCCCATCGGCGCTGGAGACACAAAGCTTGGGGGAGCCCTGCCCCACCTTGTAG 3600
QY 2695 CTCAGCGCTGTCTACCTCTCTCCAAAGAGCTGCGCGCGGGGATTTCAGAGGCGCTCACTC 2754
Db 3601 CTCAGCGCTGTCTACCTCTCTCCAAAGAGCTGCGCGCGGGGATTTCAGAGGCGCTCACTC 3660
QY 2755 AAGTGTGGGCGCAGGAGCGGCTGTGGCCCCGGGACCAAGTGCACCTTGCACCGTAAAG 2814
Db 3661 AAGTGTGGGCGCAGGAGCGGCTGTGGCCCCGGGACCAAGTGCACCTTGCACCGTAAAG 3720
QY 2815 CCCCCAGGAGCTGAGCTTCTCGCTCTGAGGCGGCTCTGA 2853
Db 3721 CCCCCAGGAGCTGAGCTTCTCGCTCTGAGGCGGCTCTGA 3759

RESULT 14

US-60-435-2152

; Sequence 2152, Application US/60230435

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

;; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
;; TITLE OF INVENTION: USES THEREOF
;; FILE REFERENCE: CL000768
;; CURRENT APPLICATION NUMBER: US/60/230,435
;; CURRENT FILING DATE: 2000-09-06
;; NUMBER OF SEQ ID NOS: 2991
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2152
;; LENGTH: 3625
;; TYPE: DNA
;; ORGANISM: HUMAN
US-60-230-435-2152

Query Match 66.9%; Score 1907.8; DB 67; Length 3625;
Best Local Similarity 77.0%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 237; Indels 546; Gaps 2;
QY 1 ATGCTTCTGCTGGGCATCTTAACCTTGGCTTTCCGCGGCGAACCCTGAGAGGCTCTGAG 60
DB 177 ATGCTTCTGCTGGGCATCTTAACCTTGGCTTTCCGCGGCGAACCCTGAGAGGCTCTGAG 236
QY 61 CCAGAGCGGAGGTAGTCTGTTCCATCCGACTGGACCCCGGACATTAACGCGCCGCGCTAC 120
DB 237 CCAGAGCGGAGGTAGTCTGTTCCATCCGACTGGACCCCGGACATTAACGCGCCGCGCTAC 296
QY 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCAGACATTT 180
DB 297 TACTGGCGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCAGACATTT 356
QY 181 CAGGAGACTTTTACCTTACACTGACGCGGATGCTGATGTTCTTGGCTCCGCGCTTCTCC 240
DB 357 CAGGAGACTTTTACCTTACACTGACGCGGATGCTGATGTTCTTGGCTCCGCGCTTCTCC 416
QY 241 ACTGAGACTTGGCGTCCCGCTCCAGGGCTCACCGGGGCTTCAGACCTGCGACGC 300
DB 417 ACTGAGACTTGGCGTCCCGCTCCAGGGCTCACCGGGGCTTCAGACCTGCGACGC 476
QY 301 TGTCTTCTATTCTGGGAGCTGAACCGCGAGCGGACTCGTTTCGCTGCTGCTGAGCCTGTGC 360
DB 477 TGTCTTCTATTCTGGGAGCTGAACCGCGAGCGGACTCGTTTCGCTGCTGAGCCTGTGC 536
QY 361 GGGGGCTTCGCGAGCCTTTGGCTACCGAGCGCGCGAGTATGTCATTAGCCCGCTGCC 420
DB 537 GGGGGCTTCGCGAGCCTTTGGCTACCGAGCGCGCGAGTATGTCATTAGCCCGCTGCC 596
QY 421 AATCTAGCGCGCGCGCGAGCAACAGCGGCGGACAGCTTCTCCAGCGCGCG 480
DB 597 AATCTAGCGCGCGCGCGAGCAACAGCGGCGGACAGCTTCTCCAGCGCGCG 656
QY 481 GGTGTTCGCGCGCGCTTCGAGAGCCCACTCTCGCTCGCGGGTGGGCTCGGGCTGG 540
DB 657 GGTGTTCGCGCGCGCTTCGAGAGCCCACTCTCGCTCGCGGGTGGGCTCGGGCTGG 716
QY 541 AACCCCGCATCTACGGGCGCTGGACCTTACAGCGCGCGGGCGGCTTCGGGGAG 600
DB 717 AACCCCGCATCTACGGGCGCTGGACCTTACAGCGCGCGGGCGGCTTCGGGGAG 776
QY 601 AGTCTAGCGCGCGAGCTTGGCGCGCGCAAGGCTTTCGCTATCCCGCGGTACGTG 660
DB 777 AGTCTAGCGCGCGAGCTTGGCGCGCGCAAGGCTTTCGCTATCCCGCGGTAGGTG 836
QY 661 GAGAGCTGTGGTTCGCGGAGAGTCAATGGTCAAGTTTCCAGCGCGCGGACCTGGAAACAT 720
DB 837 GAGAGCTGTGGTTCGCGGAGAGTCAATGGTCAAGTTTCCAGCGCGCGGACCTGGAAACAT 896
QY 721 TATCTGTAGCGCTGTGGCAACCGCGCGGACTCTACCGCATCCAGCATCTCAAC 780
DB 897 TATCTGTAGCGCTGTGGCAACCGCGCGGACTCTACCGCATCCAGCATCTCAAC 956
QY 781 CCCATCAACATCGTTGTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
DB 957 CCCATCAACATCGTTGTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 1016

QY 841 GTCACCGGCAATGGGCCCTGAGACCTGGCAACTTCTGTGCTGGCAGAGAGAGCTCAAC 900
DB 1017 GTCACCGGCAATGGGCCCTGAGACCTGGCAACTTCTGTGCTGGCAGAGAGAGCTCAAC 1076
QY 901 AAGTGTAGTACAGACACCCCGAGTACTGGACACTGCCATCTCTTTCACAGGACGAC 960
DB 1077 AAGTGTAGTACAGACACCCCGAGTACTGGACACTGCCATCTCTTTCACAGGACGAGTG 1136
QY 961 CTGTGTGAGCAGCACACCTGTGACACCCCTGGGCATGCTGATGTGGG----- 1007
DB 1137 AGTTGATCTGCGTCACTTTCACCCAGATAGTCCCGTCTTTTAGGGCATCTGGGCCAAT 1196
QY 1008 ----- 1007
DB 1197 GGAGGAGAGCCTGGCTTTCCGAAGTGTGGCCTGGCGGCCCAATCAGCGCCTCCAG 1256
QY 1008 ----- 1007
DB 1257 ATCAGGCGCGAGGGCGCGGAACCCAGGAAGTTCGCGCCCGGAGCTGCAGTTTGTGTCC 1316
QY 1008 ----- 1007
DB 1317 AAGACCGATAGGAGACGCGGTGAGATGGTGTGGAGAGGCGGGAACGGCCACCCCTA 1376
QY 1008 ----- 1007
DB 1377 TTGTATGGCGGCTGAGTCTTTCGGACACCTCCTGAGGTCCTTCAAGGGTGTAGA 1436
QY 1008 ----- 1007
DB 1437 ACTGAAGTGTATCAAGGTCAAGCTGCGCTTCTCTCGGGTAAACAGCTTGTCCCT 1496
QY 1008 ----- 1007
DB 1497 CTCGTAAACAGTCCAGGAGGATTAAGTTCCTGGCCAAAGATGATTCAGGATCATCTGA 1556
QY 1008 ----- 1007
DB 1557 AAAGTGGCTGCAGCTTACAAAGAAAGCAATTTGGGGAATCGTGAGGACAGTAGGG 1616
QY 1008 -----TACCATGTGTGACCCC 1023
DB 1617 GCTGGTTAGGATACCTACACCTTCTGGAGGAGAGGACAGATACCATGTGTGACCCC 1676
QY 1024 AAGAGAAGTGTCTGTATGAGGACGATGGCTTCATCAGCTTCCACACTGCCCCAC 1083
DB 1677 AAGAGAAGTGTCTGTATGAGGACGATGGCTTCATCAGCTTCCACACTGCCCCAC 1736
QY 1084 GAGCTGGGCCAGCTGTTCAACATGCCCATGCAATGTGAAAGTCTGTGAGGAGGTGTT 1143
DB 1737 GAGCTGGGTAAGGCTGGATAGCTCCTCTGGGGTCTTCTGGGTTTGCTGGGAGCTTT 1796
QY 1144 GGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGCACAC 1203
DB 1797 GGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGCACAC 1856
QY 1204 CCTGTGACGCTGACGTCTCCATATCAGGACTTCTGGACAGGCGGACGAGTAC 1263
DB 1857 CCTGTGACGCTGACGTCTCCATATCAGGACTTCTGGACAGGCGGACGAGTAAAG 1916
QY 1264 TGCCCTCTGACCAACCCAGACGACCTCTCCCTCCCGAGGATCTGCCGGCGGCCACAG 1323
DB 1917 CCAGAGCGGGAGGGCAATGAGCGCGCTCGAGAGGGGCTTTGCTGCCGGCGGCCAGC 1976
QY 1324 TACACCTTGAGCAGCAGTGCAGCTTGTGGCTTGTGGCTTCAAGCCCTGTCTTAC 1383
DB 1977 TACACCTTGAGCAGCAGTGCAGCTTGTGGCTTGTGGCTTCAAGCCCTGTCTTAC 2036
QY 1384 ATGACGTACTGACCAACAGCTGTGTCACCGGGAAGGCAAGGACAGATGTTGTGCCAG 1443
DB 2037 ATGACGTACTGACCAACAGCTGTGTCACCGGGAAGGCAAGGACAGATGTTGTGCCAG 2096

QY 1444 ACCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGGCGAGGCAAGCTTGGCTCAA 1503
DB 2097 ACCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGGCGAGGCAAGCTTGGCTCAA 2156
QY 1504 GGGGCTCGGTGGAGAGACACAACTCAACAGACACAGGCTGGATGTTCTGGGCC--- 1560
DB 2157 GGGGCTCGGTGGAGAGACACAACTCAACAGACACAGGCTGGATGTTCTGGAGTGC 2216
QY 1561 ----- 1560
DB 2217 GCTGGGAGCTGCTGGAGGAGGATGGATGGCTGCACACTAGCAAGGTCAGTGACTGT 2276
QY 1561 -----AAATGGGATCCCTATGGCCCTGTCTGGCGACATGTGTGGGGGTGGCAGCTG 1614
DB 2277 TACCTTTCTCCATCTGCTCAAGCGGCTAGCAGCATTTATGTGTGGGGCGTGCAGCTG 2336
QY 1615 GCCAGAGGAGCTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTG 1674
DB 2337 GCCAGAGGAGCTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTG 2396
QY 1675 AGGTGAATACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAAGAGC 1734
DB 2397 AGGTGAATACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCTCGAGGTGGGAGA 2456
QY 1735 TTCCGGGAGGAGCTGTGAGCTTTCAACGGCTTCAACAGCTTCAACACAGCACCAACCGGCTCACT 1794
DB 2457 GCAGTGGTGGCTGGCCCGAGGGAGGTGAGCTGGAACACACACACACACCGGCTCACT 2516
QY 1795 CTCGGCGTGGCATGGGTGCCAAGTACTCCGGCGGTGTCTCCCGGGACAAAGTGCAGACTC 1854
DB 2517 CTCGGCGTGGCATGGGTGCCAAGTACTCCGGCGGTGTCTCCCGGGACAAAGTGCAGACTC 2576
QY 1855 ATCTCCGAGCAANTGGCACTGGCTACTTCTATGTGTGGTGGCAACCAAGCTGGTGGAGCGC 1914
DB 2577 ATCTCCGAGCAANTGGCACTGGCTACTTCTATGTGTGGTGGCAACCAAGCTGGTGGAGCGC 2636
QY 1915 ACCTGTGTCTCTGCTGACTTCCACCTCCCTGTGTGTCTTCCAAAGCAAGTGCATCAAGCTGGC 1974
DB 2637 GGGGCTTGAGAACAAAGTAGGAGGACAGGCTTCCGGGGGAGCAAGTGCATCAAGCTGGC 2696
QY 1975 TGTGATGGGAACCTGGCTCCAAAGAGAGATTCGACAGTGTGGGTGTGGGGGAGAC 2034
DB 2697 TGTGATGGGAACCTGGCTCCAAAGAGAGATTCGACAGTGTGGGTGTGGGGGAGAC 2756
QY 2035 AATAAGAGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGTGAGTCTTGGGCGCT 2094
DB 2757 AATAAGAGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGTGAGTCTTGGGCGCT 2816
QY 2095 GTGGTGGCCATCCCGCAGGCGCTCAAGCATCGACATCGCCAGCGGGTTACAAAGG 2154
DB 2817 GAAGTCTCTGCCAGGGGCAAGAGGAGGAGGTGACATCCCGCCAGCGGGTTACAAAGG 2876
QY 2155 CTGATCGGGGATGACAACCTACCTGGCTCTGAAGAACAGCAAGCAAGTACTGCTCAAC 2214
DB 2877 CTGATCGGGGATGACAACCTACCTGGCTCTGAAGAACAGCAAGCAAGTACTGCTCAAC 2936
QY 2215 GGCATTTCTGTGTGTGGCGGTGAGAGCGGACCTGTGTGTGAAGGGCAGTCTGCTGGG 2274
DB 2937 GGCATTTCTGTGTGTGGCGGTGAGAGCGGACCTGTGTGTGAAGGGCAGTCTGCTGGG 2996
QY 2275 TACAGCGGACGGGACAGCGGTGGAGAGCCTTGCAGGCTTCCCGGCCCATCTCTGGAGCGG 2334
DB 2997 TACAGCGGACGGGACAGCGGTGGAGAGCCTTGCAGGCTTCCCGGCCCATCTCTGGAGCGG 3056
QY 2335 CTGACCGTGGAGTCTCTCTCGGTGGGAAAGATGACACCGCCCGGGTCCGCTACTCTTC 2394
DB 3057 CTGACCGTGGAGTCTCTCTCGGTGGGAAAGATGACACCGCCCGGGTCCGCTACTCTTC 3116
QY 2395 TATCTGCCAAAGAGCTCGGAGGACAGTCTCTCTATCCCAAGACCCCGGGGACCC 2454
DB 3117 TATCTGCCAAAGAGCTCGGAGGACAGTCTCTCTATCCCAAGACCCCGGGGACCC 3176
QY 2455 TCTGTCTTCACAAACAGCGTCTCTCAGCCCTCTCCACCAAGGTGGAGCAGCGGAGCAGG 2514

DB 3177 TCTGTCTTCACAAACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGACAGG 3236
QY 2515 CCCCTGCACAGCTGGTGGCTGGCAGCTGGGGGCCGTGCTCCGGAGCTCGGCGAGTGGC 2574
DB 3237 CCCCTGCACAGCTGGTGGCTGGCAGCTGGGGGCCGTGCTCCGGAGCTCGGCGAGTGGC 3296
QY 2575 CTGCAAGAACGGGGGTGGACTGTGGGGTCCCGCGGAGCGCAGCTCCCTGCTGCTGT 2634
DB 3297 CTGCAAGAACGGGGGTGGACTGTGGGGTCCCGCGGAGCGCAGCTCCCTGCTGCTGT 3356
QY 2635 GATGACAGCCCATCGGCCGTGGAGACACAAAGCCTGGGGAGCGCTCCCTCACTGGGAG 2694
DB 3357 GATGACAGCCCATCGGCCGTGGAGACACAAAGCCTGGGGAGCGCTCCCTCACTGGGAG 3416
QY 2695 CTCAGCGCTGGTCAACCTGCTCCAAAGAGCTGCGGCGGGGATTCAGAGCGCTCACTC 2754
DB 3417 CTCAGCGCTGGTCAACCTGCTCCAAAGAGCTGCGGCGGGGATTCAGAGCGCTCACTC 3476
QY 2755 AAGTGTGTGGGCCAGGAGCGCGCTGCTGGCCGGGAGCCAGTGCACCTGCAACCGCAAG 2814
DB 3477 AAGTGTGTGGGCCAGGAGCGCGCTGCTGGCCGGGAGCCAGTGCACCTGCAACCGCAAG 3536
QY 2815 CCCCAGGAGCTGGACTTCTGCTCTCTGAGCGCGTGTCTGA 2853
DB 3537 CCCCAGGAGCTGGACTTCTGCTCTCTGAGCGCGTGTCTGA 3575

RESULT 15

US-10-093-463-29

; Sequence 29, Application US/10093463

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Shenoy, Suresh

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Gusev, Vladimir

; APPLICANT: Pochart, Pascal

; APPLICANT: Zhong, Mei

; APPLICANT: Rastelli, Luca

; APPLICANT: Mezes, Peter

; APPLICANT: Smithson, Glennnda

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gerlach, Valerie

; APPLICANT: Casman, Stacie

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Tchernev, Velizar

; APPLICANT: Gangolli, Esha

; APPLICANT: Vernet, Corine

; APPLICANT: Pena, Carol

; APPLICANT: Burgess, Catherine

; APPLICANT: Liu, Xiaohong

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gorman, Linda

; APPLICANT: Spaderna, Steven

; APPLICANT: Voss, Edward

; APPLICANT: Malyankar, Uriel

; APPLICANT: Anderson, David

; APPLICANT: Patturajan, Meera

; APPLICANT: Miller, Charles

; APPLICANT: Taupier, Raymond J. Jr.

; TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic

; FILE REFERENCE: 21402-290A (Cura 590A1)

; CURRENT APPLICATION NUMBER: US/10/093,463

; CURRENT FILING DATE: 2002-06-24

; PRIOR APPLICATION NUMBER: 60/283,675

; PRIOR FILING DATE: 2001-04-14

; PRIOR APPLICATION NUMBER: 60/338,092

; PRIOR FILING DATE: 2001-12-03

; PRIOR APPLICATION NUMBER: 60/274,281

; PRIOR FILING DATE: 2001-03-08

```

1  PRIOR APPLICATION NUMBER: 60/274,101
2  PRIOR FILING DATE: 2001-03-08
3  PRIOR APPLICATION NUMBER: 60/325,681
4  PRIOR FILING DATE: 2001-09-27
5  PRIOR APPLICATION NUMBER: 60/304,354
6  PRIOR FILING DATE: 2001-07-10
7  PRIOR APPLICATION NUMBER: 60/279,995
8  PRIOR FILING DATE: 2001-03-30
9  PRIOR APPLICATION NUMBER: 60/294,899
10 PRIOR FILING DATE: 2001-05-31
11 PRIOR APPLICATION NUMBER: 60/287,424
12 PRIOR FILING DATE: 2001-04-30
13 PRIOR APPLICATION NUMBER: 60/299,027
14 PRIOR FILING DATE: 2001-06-18
15 PRIOR APPLICATION NUMBER: 60/309,198
16 PRIOR FILING DATE: 2001-07-31
17 PRIOR APPLICATION NUMBER: 60/281,194
18 PRIOR FILING DATE: 2001-04-04
19 PRIOR APPLICATION NUMBER: 60/274,194
20 PRIOR FILING DATE: 2001-03-08
21 PRIOR APPLICATION NUMBER: 60/274,849
22 PRIOR FILING DATE: 2001-03-09
23 PRIOR APPLICATION NUMBER: 60/330,380
24 PRIOR FILING DATE: 2001-10-18
25 PRIOR APPLICATION NUMBER: 60/275,235
26 PRIOR FILING DATE: 2001-03-12
27 PRIOR APPLICATION NUMBER: 60/288,342
28 PRIOR FILING DATE: 2001-05-03
29 PRIOR APPLICATION NUMBER: 60/275,578
30 PRIOR FILING DATE: 2001-03-13
31 NUMBER OF SEQ ID NOS: 370
32 SOFTWARE: PatentIn Ver. 2.1
33 SEQ ID NO 29
34 LENGTH: 2297
35 TYPE: DNA
36 ORGANISM: Homo sapiens
37 FEATURE:
38 NAME/KEY: CDS
39 LOCATION: (25)..(2290)
40 OS=10-093-463-29

```

Query Match	58.6%	Score	1671.4	DB	39	Length	2297
Best Local Similarity	79.5%	Pred. No.	0				
Matches	2267	Conservative	0	Mismatches	1	Indels	585
Gaps							
Qy	1	ATGCTTCTGCTGGGATCTTAACCTGTGCTTTCCGCGGGCGAACCGCTGGAGGCTCTGAG	60				
Db	25	ATGCTTCTGCTGGGATCTTAACCTGTGCTTTCCGCGGGCGAACCGCTGGAGGCTCTGAG	84				
Qy	61	CCAGAGCGGGAGGTAGTGGTTGCCATCCGATGGACCCGGACATTAACGGCCGCCGCTAC	120				
Db	85	CCAGAGCGGGAGGTAGTGGTTGCCATCCGATGGACCCGGACATTAACGGCCGCCGCTAC	144				
Qy	121	TACTGCGGGGTCCTCGAGGACTCGGGGATCAGGAGCTCATTTTTCAGATCACAGCATTT	180				
Db	145	TACTGCGGGGTCCTCGAGGACTCGGGGATCAGGAGCTCATTTTTCAGATCACAGCATTT	204				
Qy	181	CAGGAGGACTTTTACCTACACCTAGCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC	240				
Db	205	CAGGAGGACTTTTACCTACACCTAGCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC	264				
Qy	241	ACTGAGGATCTGGCGTCCCGCTCAGGGGCTACCGGGGCTCTTTCAGACCTTGGCAGCG	300				
Db	265	ACTGAGCATCTGGCGTCCCGCTCAGGGGCTACCGGGGCTCTTTCAGACCTTGGCAGCG	324				
Qy	301	TGCTTCTATTCTGGGACGTGAACCGGAGCTCGTTTCGCTGCTGTGAGCGCTGTGC	360				
Db	325	TGCTTCTATTCTGGGACGTGAACCGGAGCTCGTTTCGCTGCTGTGAGCGCTGTGC	384				
Qy	361	GGGGGCTCCGGAGGCTTTTGGCTACCGAGCGCCGAGTATGTCATTAGCCGCTGCC	420				
Db	385	GGGGGCTCCGGAGGCTTTTGGCTACCGAGCGCCGAGTATGTCATTAGCCGCTGCC	444				

Qy	421	AATGTAGCGCGCGCGCGCCACGCAACAGCCAGGCGCACACCTCTCTCCAGCG	GG	480
Db	445	AATGTAGCGCGCGCGCGCGCACGCAACAGCCAGGCGCACACCTCTCTCAGCGCGG	504	
Qy	481	GGTGTTCGGCGCGCGCTTCGCGAGACCCACCTCTCGCTCGCGGGTGGCCTCGGSC	1GG	540
Db	505	GGTGTTCGGCGCGCGCTTCGCGAGACCCACCTCTCGCTCGCGGGTGGCCTCGGSC	1GG	564
Qy	541	AACCCCGCCATCTTACGGGCGCTTGAGACCCCTTACAAGCCGCGGCGGGCGGTCTTCGGGGAG	600	
Db	565	AACCCCGCCATCTTACGGGCGCTTGAGACCCCTTACAAGCCGCGGCGGGCGGTCTTCGGGGAG	624	
Qy	601	AGTCGTAGCCGCGCAGCTCTGGCGCGCCAAAGGTTTCGTGTCTATCCCGGGTACGTG	660	
Db	625	AGTCGTAGCCGCGCAGCTCTGGCGCGCCAAAGGTTTCGTGTCTATCCCGGGTACGTG	684	
Qy	661	GAGACGCTGGTTCGCGGAGAGTCAATGTGTCAAGTTTCCACGCGCGGACCTTGGAACT	720	
Db	685	GAGACGCTGGTTCGCGGAGAGTCAATGTGTCAAGTTTCCACGCGCGGACCTTGGAACT	744	
Qy	721	TATCTGCTGACGCTGCTGGCAACGCGCGGAGCTCTTACCGCCATCCAGCATCTCTAAC	780	
Db	745	TATCTGCTGACGCTGCTGGCAACGCGCGGAGCTCTTACCGCCATCCAGCATCTCTAAC	804	
Qy	781	CCCATCAACATCGTTGTGTCAAGTGTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG	840	
Db	805	CCCATCAACATCGTTGTGTCAAGTGTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG	864	
Qy	841	GTACCGGCAATGGCGGCGCTGACGCTGGCGAACTTCTGTGCTTGGCAGAGAGCTCTAAC	900	
Db	865	GTACCGGCAATGGCGGCGCTGACGCTGGCGAACTTCTGTGCTTGGCAGAGAGCTCTAAC	924	
Qy	901	AAAGTCAGTGCAAGCACCCCGAGTCTGGGACACTGCCATCTCTTCAACGAGCAGGAC	960	
Db	925	AAAGTCAGTGCAAGCACCCCGAGTCTGGGACACTGCCATCTCTTCAACGAGCAGGAC	980	
Qy	961	CTGTGTGGAGCCACCACCTGTGACACCTGGGCATGGCTGATGTGGGTACCATGTGTGAC	1020	
Db	981	-----	980	
Qy	1021	CCCAAGAAAGCTGCTGTCTATTAGGACGATGGCTTCCATCAGCCTTTCACCACTGCC	1080	
Db	981	-----	980	
Qy	1081	CACGAGCTGGGCCACGTGTTCAACATGCCCATGCACATGTGAAGTCTGTGAGGAGTG	1140	
Db	981	-----	980	
Qy	1141	TTTGGGAAGCTCGGAGCCAAACCACATGATGTCCCGACCTTCATCCAGATCGACCGTGCC	1200	
Db	981	-----	980	
Qy	1201	AAOCCCTGGTGCAGCTGCAGTGTGCCATCATCACCAGTTCCTTGGACAGCGGCACGGT	1260	
Db	981	-----	980	
Qy	1261	GACTGCTCTTGACCAACCCAGAACGCCATCTCCCTGCCCGAGGATCTCCCGGGCGCC	1320	
Db	981	-----	980	
Qy	1321	AGCTACACCTGAGCCAGCAGTGGCAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCT	1380	
Db	981	-----	980	
Qy	1381	TACATGCAGTACTGCACCAAGCTGTGGTGCACCGGGAAGGCCAAGGACAGATGCTGTGC	1440	
Db	981	-----	980	
Qy	1441	CAGACCCGCCACTTCCCTGGCGCGATGSCACCACTGTGGCAGGGCAAGCTCTGCCTC	1500	
Db	981	-----	980	
Qy	1501	AAAGGGCCTCGCTGGAGAGACACAACCTCAACAAGCACAGGTTGGATGTCTCTGGGCC	1560	

```
Db 981 -----GCTGGATGTTCTCTGGCC 999
QY 1561 AAATGGATCCCTATGGCCCTCTGTCGGGCACATGTGTTGGGGCGTGCACCTGGGCCAGG 1620
Db 1000 AAATGGATCCCTATGGCCCTCTGTCGGGCACATGTGTTGGGGCGTGCACCTGGGCCAGG 1059
QY 1621 AGGCACTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGCAGGGTG 1680
Db 1060 AGGCACTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGCAGGGTG 1119
QY 1681 AAATACCGATTCGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCCGAAAGAGTTCGCG 1740
Db 1120 AAATACCGATTCGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCCGAAAGAGTTCGCG 1179
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAAACAGCAGCACCAACCGGCTCACTCTCGCC 1800
Db 1180 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAAACAGCAGCACCAACCGGCTCACTCTCGCC 1239
QY 1801 GTGGCATGGGTGCCAAGTACTTCGGGCTGTCTCCCGGGACAAAGTGCAGCTCATCTGC 1860
Db 1240 GTGGCATGGGTGCCAAGTACTTCGGGCTGTCTCCCGGGACAAAGTGCAGCTCATCTGC 1299
QY 1861 CGAGCCAATGGCACTGGCTACTTCTATGTGTGGCACCAAGTGTGGAGCGCACGCTG 1920
Db 1300 CGAGCCAATGGCACTGGCTACTTCTATGTGTGGCACCAAGTGTGGAGCGCACGCTG 1359
QY 1921 TGCTCTCTGACTCCACCTCCCTCTGTGTCCAAAGCCATGCATGGCTCAATTTCTGGTG 1980
Db 1360 TGCTCTCTGACTCCACCTCCCTCTGTGTCCAAAGCCATGCATGGCTCAATTTCTGGTG 1419
QY 1981 GGCACACTGGCTCCAAAGAGAGATTCGACAGTGTGGGTGTGTGGGGAGAGCAATAAG 2040
Db 1420 GGCACACTGGCTCCAAAGAGAGATTCGACAGTGTGGGTGTGTGGGGAGAGCAATAAG 1479
QY 2041 AGCTGCAAGAAGTGAAGTCTTCAACAGCCCATGCATGGCTCAATTTCTGGTG 2100
Db 1480 AGCTGCAAGAAGTGAAGTCTTCAACAGCCCATGCATGGCTCAATTTCTGGTG 1539
QY 2101 GCCATCCCGCAGGCGCTTCAGCATCGCATCCCGAGCGGTTTCAAAAGGCTGATC 2160
Db 1540 GCCATCCCGCAGGCGCTTCAGCATCGCATCCCGAGCGGTTTCAAAAGGCTGATC 1599
QY 2161 GGGGATGACAACCTACCTGGCTCTGAAGAACAGCCAGTACCTGCTCAAGGGGAT 2220
Db 1600 GGGGATGACAACCTACCTGGCTCTGAAGAACAGCCAGTACCTGCTCAAGGGGAT 1659
QY 2221 TTCTGTGTGTCGGCGTGGAGCGGACCTGTGTGTAAGGGAGTCTGCTGCGGTACAGC 2280
Db 1660 TTCTGTGTGTCGGCGTGGAGCGGACCTGTGTGTAAGGGAGTCTGCTGCGGTACAGC 1719
QY 2281 GGCACGGGCACAGCGTGGAGAGCTTGCAGGCTTCCCGGCCATCTCTGGAGCGCTGACC 2340
Db 1720 GGCACGGGCACAGCGTGGAGAGCTTGCAGGCTTCCCGGCCATCTCTGGAGCGCTGACC 1779
QY 2341 GTGGAGTCTCTTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 1780 GTGGAGTCTCTTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 1839
QY 2401 CCCAAGAGCTCTGGAGGACAGTCTCTCATCCAGGACCCCGGGACCCCTCTGTC 2460
Db 1840 CCCAAGAGCTCTGGAGGACAGTCTCTCATCCAGGACCCCGGGACCCCTCTGTC 1899
QY 2461 TTGCACAACAGCTCTCAAGCTCTCCAAACAGGTGGAGCGCGGAGACAGGCCCT 2520
Db 1900 TTGCACAACAGCTCTCAAGCTCTCCAAACAGGTGGAGCGCGGAGACAGGCCCT 1959
QY 2521 GCAGCTGGGTGCTGGCAGCTGGGGCGGCTGCTCCGGAGCTGGGGAGTGGCTGCAG 2580
Db 1960 GCAGCTGGGTGCTGGCAGCTGGGGCGGCTGCTCCGGAGCTGGGGAGTGGCTGCAG 2019
QY 2581 AAGCGGCGGTGACTGTGGGGCTCCGGCGGAGCGAGCTGCTGCTGTGATGCA 2640
```

```
Db 2020 AAGCGGCGGTGGACTGCGGGGCTCCGCCGGCAGCGCTCCCTGCTGTGATGCA 2079
QY 2641 GCCCATCGGCCCTGGAGACACAAAGCTTGGGGAGCCCTGCCCACTGGGAGTTCAGC 2700
Db 2080 GCCCATCGGCCCTGGAGACACAAAGCTTGGGGAGCCCTGCCCACTGGGAGTTCAGC 2139
QY 2701 GCTGTGTACCCCTGCTCCAAAGAGTGGCGCGGGGATTTTCAGAGCGCTCACYCAATGT 2760
Db 2140 GCTGTGTACCCCTGCTCCAAAGAGTGGCGCGGGGATTTTCAGAGCGCTCACYCAATGT 2199
QY 2761 GTGGGCCAGGGAGCGGCTGCTGGCCGGGAGCAGTGCACCTTGCACCCGAGCCAG 2820
Db 2200 GTGGGCCAGGGAGCGGCTGCTGGCCGGGAGCAGTGCACCTTGCACCCGAGCCAG 2259
QY 2821 GAGCTGGACTTCTGGCTCTGAGGCCGTGCTGA 2853
Db 2260 GAGCTGGACTTCTGGCTCTGAGGCCGTGCTGA 2292
```

Search completed: May 16, 2003, 00:28:37
Job time : 5635 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 22:28:13 ; Search time 1084 Seconds
(without alignments)
12487.358 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtcttctgtggcctcct.....gcgtcttgagcggtgtcgtga 2853

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 614469 seqs, 2372291704 residues

Total number of hits satisfying chosen parameters: 1228938

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
10: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2851.4	99.9	2853	8	US-10-391-364-76
2	2851.4	99.9	2853	8	US-10-391-364-78
3	2851.4	99.9	2853	8	US-10-170-235-10538
4	2848.2	99.8	2853	11	US-60-453-135-7278
5	2848.2	99.8	2853	11	US-60-453-050-7278
6	2847.6	99.8	2853	9	US-10-311-035-32
7	2299.4	80.6	5714	9	US-10-144-771-12354
8	664.8	23.3	3933	10	US-60-455-444-3860
9	664.4	23.3	3711	6	US-09-949-002-108
10	664.4	23.3	4016	6	US-09-724-676A-18640
11	664.4	23.3	4016	6	US-09-724-676A-18640
12	662.8	23.2	2664	6	US-09-949-002-248
13	662.8	23.2	3933	8	US-10-170-235-34404
14	636	22.3	2856	9	US-10-144-771-20257
15	629.8	22.1	4309	1	PCT-US02-24567-57
16	629.8	22.1	4309	9	US-10-210-120-57
17	629.8	22.1	4439	8	US-10-170-235-18710
18	629.8	22.1	4459	9	US-10-159-563-192
19	629.8	22.1	4459	9	US-10-159-563-308
20	629	22.0	4639	10	US-60-455-444-3745
21	629	22.0	4639	11	US-60-452-680-11289
22	629	22.0	4639	11	US-60-453-135-7150

23	629	22.0	4639	11	US-60-453-050-7150	Sequence 7150, Ap
24	628.2	22.0	2853	1	PCT-US02-41225A-18	Sequence 18, Appl
25	628.2	22.0	3430	1	PCT-US02-41225A-17	Sequence 17, Appl
26	628.2	22.0	4471	11	US-60-440-068-133	Sequence 133, App
27	612.8	21.5	4878	9	US-10-152-319A-1840	Sequence 1840, Ap
28	612.8	21.5	4878	9	US-10-301-856-1040	Sequence 1040, Ap
29	609.6	21.5	4878	9	US-10-338-044-2196	Sequence 2196, Ap
30	609.6	21.4	3172	6	US-09-724-676-32233	Sequence 32233, A
31	609.6	21.4	3172	6	US-09-724-676A-32233	Sequence 32233, A
32	585.6	20.5	4305	10	US-60-455-444-1128	Sequence 1128, Ap
33	585.6	20.5	4305	11	US-60-452-680-3792	Sequence 3792, Ap
34	585.6	20.5	4305	11	US-60-453-135-2298	Sequence 2298, Ap
35	585.6	20.5	4305	11	US-60-453-050-2298	Sequence 2298, Ap
36	583.2	20.4	2514	1	PCT-US02-41225A-20	Sequence 20, Appl
37	583.2	20.4	3593	6	US-09-949-002-278	Sequence 278, App
38	583.2	20.4	4192	5	US-09-634-2870-1	Sequence 1, Appl
39	583.2	20.4	4301	1	PCT-US02-41225A-19	Sequence 19, Appl
40	583.2	20.4	4305	8	US-10-170-235-30070	Sequence 30070, A
41	583.2	20.4	4307	6	US-09-949-002-94	Sequence 94, Appl
42	583.2	20.4	4307	8	US-10-358-283-31	Sequence 31, Appl
43	581.6	20.4	4407	8	US-10-015-392A-316	Sequence 316, App
44	581.6	20.4	4407	8	US-10-013-912A-316	Sequence 316, App
45	581.6	20.4	4407	8	US-10-015-394A-316	Sequence 316, App

ALIGNMENTS

RESULT 1

US-10-391-364-76
; Sequence 76, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroli, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MPI03-0190NMIM
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

NAME/KEY: CDS
LOCATION: (1)...(2853)
US-10-391-364-76

Query Match 99.9%; Score 2851.4; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTTTCTGCTGGGATCCTAACCTTGGCTTTCGGCGGGCAACCGCTGGAGGCTCTGAG 60
Db 1 ATCTTTCTGCTGGGATCCTAACCTTGGCTTTCGGCGGGCAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGGACCGCGACATTAACGGCGCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGGACCGCGACATTAACGGCGCGCTAC 120
QY 121 TACTGGGGGTTCCAGAGACTCCGGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180
Db 121 TACTGGGGGTTCCAGAGACTCCGGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180
QY 181 CAGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTAGCATCTGGGCGTCCCGCTCCAGGGGCTCAACGGGGCTTTCAGACCTGGGACGC 300
Db 241 ACTAGCATCTGGGCGTCCCGCTCCAGGGGCTCAACGGGGCTTTCAGACCTGGGACGC 300
QY 301 TGCTTTATTTCTGGGACGTGAACGGCGGAGCTCGTTCGCTGTGAGCCCTGTGC 360
Db 301 TGCTTTATTTCTGGGACGTGAACGGCGGAGCTCGTTCGCTGTGAGCCCTGTGC 360
QY 361 GGGGGCTCCCGGAGCTTTGGCTACCGAGGCGCGAGTATGCTATTAGCCGCTGCC 420
Db 361 GGGGGCTCCCGGAGCTTTGGCTACCGAGGCGCGAGTATGCTATTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGGCGAGCGAAGACGAGCGGCGCACACCTTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGGGGCGAGCGAAGACGAGCGGCGCACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGCGGGCTTCCGAGACCCACCTCTCGCTGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCCGGCGGGCTTCCGAGACCCACCTCTCGCTGGGGTGGCTCGGGCTGG 540
QY 541 AACCCGCCATCTACGGGCGCTTGAACCTTACAGCGCGGCGGGCGGCTTCGGGGAG 600
Db 541 AACCCGCCATCTACGGGCGCTTGAACCTTACAGCGCGGCGGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGAGGTCTGGGCGGCGCAAGCTTTGCTGTATCCCGCGGTACGTG 660
Db 601 AGTCGTAGCGCGCGAGGTCTGGGCGGCGCAAGCTTTGCTGTATCCCGCGGTACGTG 660
QY 661 GAGACGTGTTGGTTCGGGAGAGTCAATGGTCAAGTTCCACGGCGGAGCTTGAACAT 720
Db 661 GAGACGTGTTGGTTCGGGAGAGTCAATGGTCAAGTTCCACGGCGGAGCTTGAACAT 720
QY 721 TATCTGTGACGTGTGGACGCGGCGGAGCTTACCGCATCCAGACTCTCTCAAC 780
Db 721 TATCTGTGACGTGTGGACGCGGCGGAGCTTACCGCATCCAGACTCTCTCAAC 780
QY 781 CCCATCAACATCTGTTGGTCAAGGTCGCTTCTTAGAGATCGTACCTCCGGGCCAAG 840
Db 781 CCCATCAACATCTGTTGGTCAAGGTCGCTTCTTAGAGATCGTACCTCCGGGCCAAG 840
QY 841 GTCACCGCAATGGGCGCTGAGCTCGCAACTTCTGCGCTGGAGAGAGAGCTGAAC 900
Db 841 GTCACCGCAATGGGCGCTGAGCTCGCAACTTCTGCGCTGGAGAGAGAGCTGAAC 900
QY 901 AAAGTGAGTGACAGACCCCGAGTACTGGGACACTGCCATCTCTTCCAGGCGAGGAC 960
Db 901 AAAGTGAGTGACAGACCCCGAGTACTGGGACACTGCCATCTCTTCCAGGCGAGGAC 960
QY 961 CTGTGTGGAGCCACCCTGTGACACCCCTGGGATGGCTGATGTGGGTACCATGTGTGAC 1020

Db 961 CTGTGTGGAGCCACCCTGTGACACCCCTGGCATGGCTGATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGTCTATTGAGGAGATGGGCTTCCATCAGCTTTCACCACTGCC 1080
Db 1021 CCCAAGAGAGCTGCTCTGTCTATTGAGGAGATGGGCTTCCATCAGCTTTCACCACTGCC 1080
QY 1081 CACGAGCTGGGCGGACGTGTTCAACATCCCCCATGACAAATGTAAAGTCTGTGAGGAGTG 1140
Db 1081 CACGAGCTGGGCGGACGTGTTCAACATCCCCCATGACAAATGTAAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACCCACATGATTCGCCGACCTCATCCAGATCGACCTGCC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACCCACATGATTCGCCGACCTCATCCAGATCGACCTGCC 1200
QY 1201 AACCCCTGGTCAAGTGTGCTGCTCATCACCGACTTCTCTGGACAGCGGCGACGGT 1260
Db 1201 AACCCCTGGTCAAGTGTGCTGCTCATCACCGACTTCTCTGGACAGCGGCGACGGT 1260
QY 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCTCCCTGGCCGAGGATCTGCCGGTGC 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCTCCCTGGCCGAGGATCTGCCGGTGC 1320
QY 1321 AGCTACACCTTGAGCCAGCTGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCTCT 1380
Db 1321 AGCTACACCTTGAGCCAGCTGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCTCT 1380
QY 1381 TACTGACAGTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAAGGACAGATGGTGTGC 1440
Db 1381 TACTGACAGTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAAGGACAGATGGTGTGC 1440
QY 1441 CAGACCCGCCACTTCCCTGGGCGGATGGACACAGCTGTGGCGAGGCGCAAGCTCTGCTC 1500
Db 1441 CAGACCCGCCACTTCCCTGGGCGGATGGACACAGCTGTGGCGAGGCGCAAGCTCTGCTC 1500
QY 1501 AAAGGGGCTCGCTGGAGAGACAACTCAACAGCACAGGGTGGATGGTCTCTGGGCG 1560
Db 1501 AAAGGGGCTCGCTGGAGAGACAACTCAACAGCACAGGGTGGATGGTCTCTGGGCG 1560
QY 1561 AAATGGGATCCTATGGCCCTGTCTCGGCACATGTGTGGGGGCGGTGCAGCTGGCAGG 1620
Db 1561 AAATGGGATCCTATGGCCCTGTCTCGGCACATGTGTGGGGGCGGTGCAGCTGGCAGG 1620
QY 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGGAGTACTCGGAGGAGTCAAGGTG 1680
Db 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGGAGTACTCGGAGGAGTCAAGGTG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAGAGCTTCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAGAGCTTCGG 1740
QY 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTAAACACAGCACCAACCGGCTCACTCTGCC 1800
Db 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTAAACACAGCACCAACCGGCTCACTCTGCC 1800
QY 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGACAGTGCAGCTCATCTGC 1860
Db 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGACAGTGCAGCTCATCTGC 1860
QY 1861 CGAGCAATGGCAGTGGCTTCTATGTGTGGCACCACCAAGGTGGTGGACGGCACGCTG 1920
Db 1861 CGAGCAATGGCAGTGGCTTCTATGTGTGGCACCACCAAGGTGGTGGACGGCACGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGGAACTGGGCTCCAAAGAGATTCACAAAGTGTGGGTGTGTGGGGGAGACAAAG 2040
Db 1981 GGGAACTGGGCTCCAAAGAGATTCACAAAGTGTGGGTGTGTGGGGGAGACAAAG 2040
QY 2041 AGCTGCAAGAGGCTGACTGGACTCTTCAACAGCCCATGCTACAAATTCGTGGTG 2100

Db 2041 AGCTGCAAGAAGGTGACTGGACTCTTACCAAGCCCATGATGGCTACAAATTTCTGGTGTG 2100
QY 2101 GCCATCCCGCAGGCGCTCAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGCAGGCGCTCAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2160
QY 2161 GGGGATGACAACTACTGGCTCTGAAGAACAGCAAGCAAGTACTTGTCTCAACGGGCAT 2220
Db 2161 GGGGATGACAACTACTGGCTCTGAAGAACAGCAAGCAAGTACTTGTCTCAACGGGCAT 2220
QY 2221 TTCGTGGTGTCCGGGTGGAGCGGACCTGTGTGAAGGCAGTCTGCTCCGCTATCTG 2280
Db 2221 TTCGTGGTGTCCGGGTGGAGCGGACCTGTGTGAAGGCAGTCTGCTCCGCTATCTG 2280
QY 2281 GGCAGCGGCACAGCGGTGGAGCGCTGAGGCTTCCCGGCCCATCTGGAGCCGCTGACC 2340
Db 2281 GGCAGCGGCACAGCGGTGGAGCGCTGAGGCTTCCCGGCCCATCTGGAGCCGCTGACC 2340
QY 2341 GTGGAGGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTCTATCTG 2400
Db 2341 GTGGAGGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGACCCCGGGGACCCCTCTGTG 2460
Db 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGACCCCGGGGACCCCTCTGTG 2460
QY 2461 TTGCACACAGCGTCTCAGGCTCTCCAAACAGGTGGAGCAGCGGAGCAGCCGCT 2520
Db 2461 TTGCACACAGCGTCTCAGGCTCTCCAAACAGGTGGAGCAGCGGAGCAGCCGCT 2520
QY 2521 GCACGCTGGGTGGCAGTGGGGCCGTCTCCGGAGCTGCGGAGTGGGCTGCGAG 2580
Db 2521 GCACGCTGGGTGGCAGTGGGGCCGTCTCCGGAGCTGCGGAGTGGGCTGCGAG 2580
QY 2581 AAGCGGGGGTGGTGTCTGGGGTCTCCGGGGGACCGACCGTCTCTGCTGTGATGCA 2640
Db 2581 AAGCGGGGGTGGTGTCTGGGGTCTCCGGGGGACCGACCGTCTCTGCTGTGATGCA 2640
QY 2641 GCCATCGCGCGTGAGACACAAAGCCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
Db 2641 GCCATCGCGCGTGAGACACAAAGCCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCACTCTCCAAAGCTGCGCGCGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGGTCACTCTCCAAAGCTGCGCGCGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGAGGAGCGGCTGTGCGCGGAGCAGTGTGCACTTGCACCGCAAGCCCGAG 2820
Db 2761 GTGGGCGAGGAGCGGCTGTGCGCGGAGCAGTGTGCACTTGCACCGCAAGCCCGAG 2820
QY 2821 GAGCTGGACTTCTGCTGCTGAGGCGGTGCTGA 2853
Db 2821 GAGCTGGACTTCTGCTGCTGAGGCGGTGCTGA 2853

RESULT 2
US-10-391-364-78
; Sequence 78, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MP103-0190NMIM
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370

; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2853)
US-10-391-364-78

Query Match 99.9%; Score 2851.4; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTTAACCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCTTAACCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60

QY 61 CCAGAGCGGAGGTAGTCGTTCCATCCGACTGACCCGCGACATTAAACGCGCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTCGTTCCATCCGACTGACCCGCGACATTAAACGCGCGCTAC 120

QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180

QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCCTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCCTGGCTCCCGCTTCTCC 240

QY 241 ACTGAGCATCTGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGCGACGC 300
Db 241 ACTGAGCATCTGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGCGACGC 300

QY 301 TGTCTTCTATTCTGGGACGTGAACCGCGAGCTGCTTCGCTGCTGTGAGCCTGTGC 360
Db 301 TGTCTTCTATTCTGGGACGTGAACCGCGAGCTGCTTCGCTGCTGTGAGCCTGTGC 360

QY 361 GGGGGGCTCCCGAGGCTTTGGCTACCGGGGCTCAGTATGTCATTAGCCGCTGCC 420
Db 361 GGGGGGCTCCCGAGGCTTTGGCTACCGGGGCTCAGTATGTCATTAGCCGCTGCC 420

QY 421 AATGCTAGCGCGCGCGCGCAGCAACAGCGCGACACCTTCTTCAGAGCGCGG 480
Db 421 AATGCTAGCGCGCGCGCGCAGCAACAGCGCGACACCTTCTTCAGAGCGCGG 480

QY 481 GGTGTTCCGGGCGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCCGGGCGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540

QY 541 AACCCCGCCATCTTACGGGCGCTTACAGCCCTTACAGCCGCGGGGCGGCTTCGGGGAG 600
Db 541 AACCCCGCCATCTTACGGGCGCTTACAGCCCTTACAGCCGCGGGGCGGCTTCGGGGAG 600

Db 541 AACCCGCCATCCTACGGGCCCTGGACCCCTTACAAGCCGCGGGGGGCGCTTCGGGGAG 600
QY 601 AGTCGTAGCCGGCGAGGTCTGGCGCGCCAAAGCGTTTCGTGTCTATPCCCGCGGTAGCTG 660
Db 601 AGTCGTAGCCGGCGAGGTCTGGCGCGCCAAAGCGTTTCGTGTCTATATCCCGCGGTAGCTG 660
QY 661 GAGACGCTGGTGGTCCGCGGAGGTCAATGGTCAAGTTTCCACGGCGGGACCTTGGACAT 720
Db 661 GAGACGCTGGTGGTCCGCGGAGGTCAATGGTCAAGTTTCCACGGCGGGACCTTGGACAT 720
QY 721 TATCTGTGACGTCTGGCAAGCGCGCGGCGACTCTACCGCCATCCACGATCCTCAAC 780
Db 721 TATCTGTGACGTCTGGCAAGCGCGCGGCGACTCTACCGCCATCCACGATCCTCAAC 780
QY 781 CCCATCAACATCGTTGTGGTCAAGTCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
Db 781 CCCATCAACATCGTTGTGGTCAAGTCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
QY 841 GTCACCGCAATGGCCCTGACGCTGCGCAACTTCTGTCCCTGGCAGAGAAGCTGAAC 900
Db 841 GTCACCGCAATGGCCCTGACGCTGCGCAACTTCTGTCCCTGGCAGAGAAGCTGAAC 900
QY 901 AAAGTGAGTGACAAAGCACCCCGAGTACTGGGACACTGCCATCCTCTTCACAGGCGAGGAC 960
Db 901 AAAGTGAGTGACAAAGCACCCCGAGTACTGGGACACTGCCATCCTCTTCACAGGCGAGGAC 960
QY 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGGCACTGGCTGATGGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGGCACTGGCTGATGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGCTGCTGTCTATGTAGGACGATGGCTTCCATCAGCCTTACCACCTGCC 1080
Db 1021 CCCAAGAGAAGCTGCTGTCTATGTAGGACGATGGCTTCCATCAGCCTTACCACCTGCC 1080
QY 1081 CACGAGCTGGGCGACGTTTCAACATGCCCCATGACAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CACGAGCTGGGCGACGTTTCAACATGCCCCATGACAATGTGAAGTCTGTGAGGAGGTG 1140
QY 1141 TTTGGGAAGCTCGAGCCACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGC 1200
Db 1141 TTTGGGAAGCTCGAGCCACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGC 1200
QY 1201 AACCCCTGGTCAGCTGAGTGCATCATCACCAGCTTCTGTGACAGCGGGCACGT 1260
Db 1201 AACCCCTGGTCAGCTGAGTGCATCATCACCAGCTTCTGTGACAGCGGGCACGT 1260
QY 1261 GACTGCTCTGGACCAACCCAGCCCATCTCCCTGCCCCGAGGATCTGCCGGCGCC 1320
Db 1261 GACTGCTCTGGACCAACCCAGCCCATCTCCCTGCCCCGAGGATCTGCCGGCGCC 1320
QY 1321 AGCTACACCTGAGCAGAGTGCAGTGGCTTTTGGCGTGGGCTCCAAGCCCTGTCT 1380
Db 1321 AGCTACACCTGAGCAGAGTGCAGTGGCTTTTGGCGTGGGCTCCAAGCCCTGTCT 1380
QY 1381 TACATGAGTACGACCAAGCTGTGGTGCACCGGGAAGGCCAAGGAGACATGTGTGC 1440
Db 1381 TACATGAGTACGACCAAGCTGTGGTGCACCGGGAAGGCCAAGGAGACATGTGTGC 1440
QY 1441 CAGACCCGCCATTCCTCGGCGCGATGGCACAGCTGTGGCGAGGCAAGCTCTGCCCTC 1500
Db 1441 CAGACCCGCCATTCCTCGGCGCGATGGCACAGCTGTGGCGAGGCAAGCTCTGCCCTC 1500
QY 1501 AAAGGGCCCTGCGTGGAGAGACACACCTCAACAGCAGGAGGCTGATGGTTCCTGGGCG 1560
Db 1501 AAAGGGCCCTGCGTGGAGAGACACACCTCAACAGCAGGAGGCTGATGGTTCCTGGGCG 1560
QY 1561 AAATGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGGCGGTGACGTGGCCAGG 1620
Db 1561 AAATGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGGCGGTGACGTGGCCAGG 1620
QY 1621 AGGCAATGACCAACCCCAACCCCTGCGCAACAGCGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
Db 1621 AGGCAATGACCAACCCCAACCCCTGCGCAACAGCGGGCAAGTACTGCGAGGAGTGAAGGTG 1680

QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAGAGCTTCCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAGAGCTTCCGG 1740
QY 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTACACACAGCAGCAGCCGCTCAGTCTGCC 1800
Db 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTACACACAGCAGCAGCCGCTCAGTCTGCC 1800
QY 1801 GTGGCATGGTCCCAAGTACTCCCGCGGTGTCTCCCGGGACAGTCAAGCTCATCTGC 1860
Db 1801 GTGGCATGGTCCCAAGTACTCCCGCGGTGTCTCCCGGGACAGTCAAGCTCATCTGC 1860
QY 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGCACCAAGGTGGTGAGCGGACCGCTG 1920
Db 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGCACCAAGGTGGTGAGCGGACCGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCCGTCTGTGTCAAGCAAGTGCATCAAGGCTGGGTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCCGTCTGTGTCAAGCAAGTGCATCAAGGCTGGGTGTGAT 1980
QY 1981 GGGAACTGGGCTCCCAAGAAGATTCACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGGCTCCCAAGAAGATTCACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTCCCAAGAAGTGTGACTTCTTCAACAGCCCATGCATGGCTACAATTTCTGTG 2100
Db 2041 AGCTCCCAAGAAGTGTGACTTCTTCAACAGCCCATGCATGGCTACAATTTCTGTG 2100
QY 2101 GCCATCCCGCAGGGCCCTCAAGNATCGACATCCGCCAGCGGTTTCAAGGGCTGATC 2160
Db 2101 GCCATCCCGCAGGGCCCTCAAGNATCGACATCCGCCAGCGGTTTCAAGGGCTGATC 2160
QY 2161 GGGGATGACAACTACTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGCAT 2220
Db 2161 GGGGATGACAACTACTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGCAT 2220
QY 2221 TTCTGTGTGTGGGGTGGAGCGGACCTGTGTGAAGGCGAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCTGTGTGTGGGGTGGAGCGGACCTGTGTGAAGGCGAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCAGGCGCAGCGGTGGAGAGCTTCAGGCTTCCCGGCCCATCTGGAGCCCTGACC 2340
Db 2281 GGCAGGCGCAGCGGTGGAGAGCTTCAGGCTTCCCGGCCCATCTGGAGCCCTGACC 2340
QY 2341 GTGGAGGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2341 GTGGAGGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCCCTCTGTC 2460
Db 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCCCTCTGTC 2460
QY 2461 TTGCACACACGCTCTCAGGCTTCTCAACAGGTGAGAGCGCGGAGCAGAGGCCCT 2520
Db 2461 TTGCACACACGCTCTCAGGCTTCTCAACAGGTGAGAGCGCGGAGCAGAGGCCCT 2520
QY 2521 GCAGCTGGGTGGTGGCAGCTGGGGCGGTCTCCCGGAGCTCGCGAGCTGGGCTGCGAG 2580
Db 2521 GCAGCTGGGTGGTGGCAGCTGGGGCGGTCTCCCGGAGCTCGCGAGCTGGGCTGCGAG 2580
QY 2581 AAGCGGGGTGGACTGTCCGGGGTCCCGGGGAGCGGCAAGGCTGCTGTGTGATGCA 2640
Db 2581 AAGCGGGGTGGACTGTCCGGGGTCCCGGGGAGCGGCAAGGCTGCTGTGTGATGCA 2640
QY 2641 GCCCATGGCCCGTGGAGACACAGCTTGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
Db 2641 GCCCATGGCCCGTGGAGACACAGCTTGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGCTCCAAAGAGTGTGGGGTCCCGGGGAGGATTTTCAGAGGCGCTCAAGTGT 2760
Db 2701 GCCTGGTCAACCTGCTCCAAAGAGTGTGGGGTCCCGGGGAGGATTTTCAGAGGCGCTCAAGTGT 2760

QY 2761 GTGGCCACGAGCGGCTGCTGGCCGGGACCAAGTGCACCTTGACCGCAAGCCCGAG 2820
DB 2761 GTGGCCACGAGCGGCTGCTGGCCGGGACCAAGTGCACCTTGACCGCAAGCCCGAG 2820
QY 2821 GAGCTGGACTTCTGGCTGCTGAGCGCGTGTGA 2853
DB 2821 GAGCTGGACTTCTGGCTGCTGAGCGCGTGTGA 2853

RESULT 3
US-10-170-235-10538
; Sequence 10538, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CLO01380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 10538
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-10538

Query Match 99.98; Score 2851.4; DB 8; Length 2853;
Best Local Similarity 100.08; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCCTAACCTGCTTTCGGCGGGGCAACCGGTGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGATCCTAACCTGCTTTCGGCGGGGCAACCGGTGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGTGTCTGCCATCCGACTGGACCGGACATTAAGCGCGCGGTAC 120
DB 61 CCAGAGCGGGAGTGTCTGCCATCCGACTGGACCGGACATTAAGCGCGCGGTAC 120
QY 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGCTATTTTCAGATCACAGCATTT 180
DB 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGCTATTTTCAGATCACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACTGACGCGGATGCTAGTTGCTTGGCTCCCGCTTCC 240
DB 181 CAGGAGGACTTTTACCTACACTGACGCGGATGCTAGTTGCTTGGCTCCCGCTTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTACCGGGGGCTCTTACAGCCTCGGAGC 300
DB 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTACCGGGGGCTCTTACAGCCTCGGAGC 300
QY 301 TCGTTCTATTCTGGGAGCTGAACCGCGAGCGGACTCGTTCTGCTGTGCTGTGTC 360
DB 301 TCGTTCTATTCTGGGAGCTGAACCGCGAGCGGACTCGTTCTGCTGTGCTGTGTC 360
QY 361 GGGGGCTCCCGGAGCCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCCGCTGCC 420
DB 361 GGGGGCTCCCGGAGCCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGCGGCGGACGCGCAACAGCGGCGGACACCTTCTCCAGGCGCG 480
DB 421 AATGCTAGCGCGCGGCGGCGGACGCGCAACAGCGGCGGACACCTTCTCCAGGCGCG 480
QY 481 GGTGTTCCGGGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGTGG 540
DB 481 GGTGTTCCGGGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGTGG 540
QY 541 AACCCCGGCATCCTACGGGCCCTTGACCCCTTACAGCGCGCGGCGGCTTCGGGGAG 600
DB 541 AACCCCGGCATCCTACGGGCCCTTGACCCCTTACAGCGCGCGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGGAGGTCTGGGCGCGCGCAAGCGCTTTCTGTCTATCCCGGTTACGTG 660
DB 601 AGTCGTAGCGCGGAGGTCTGGGCGCGCGCAAGCGCTTTCTGTCTATCCCGGTTACGTG 660

QY 661 GAGACGCTGCTGCTCGCGGACGAGTCAATGGTCAAGTTCCAGCGCGGACCTGAAACAT 720
DB 661 GAGACGCTGCTGCTCGCGGACGAGTCAATGGTCAAGTTCCAGCGCGGACCTGAAACAT 720
QY 721 TATCTGCTGACGCTGCTGGCAAGCGCGGCGGACTCTACCGCATCCCGAGCATCCCAAC 780
DB 721 TATCTGCTGACGCTGCTGGCAAGCGCGGCGGACTCTACCGCATCCCGAGCATCCCAAC 780
QY 781 CCCATCAACATCGTTGCTGTCAGAGTGTCTTCTTAGAGATCGTGACTCCGGGCCCCAAG 840
DB 781 CCCATCAACATCGTTGCTGTCAGAGTGTCTTCTTAGAGATCGTGACTCCGGGCCCCAAG 840
QY 841 GTCACCGGCAATCGCGCCCTGACCGCTGCGCAACTTCTGTGCTGGCAGAGAGCTGAAC 900
DB 841 GTCACCGGCAATCGCGCCCTGACCGCTGCGCAACTTCTGTGCTGGCAGAGAGCTGAAC 900
QY 901 AAAGTGAGTGACAAAGCAACCCCGGACTACTGGGACACTGCCATCCTTTCACAGGCGAGAC 960
DB 901 AAAGTGAGTGACAAAGCAACCCCGGACTACTGGGACACTGCCATCCTTTCACAGGCGAGAC 960
QY 961 CTGCTGGAGCCACCACCTGTGACACCCCTGGGCATGGCTGATGGGTACCATGTGTGAC 1020
DB 961 CTGCTGGAGCCACCACCTGTGACACCCCTGGGCATGGCTGATGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGACGATGGGCTTCCATCAGCCTTCACCACTGCC 1080
DB 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGACGATGGGCTTCCATCAGCCTTCACCACTGCC 1080
QY 1081 CACGAGTGGGCGCACGCTGTGCAATGCCCATGCGGACACTGTGAAGTCTGTGAGAGTGTG 1140
DB 1081 CACGAGTGGGCGCACGCTGTGCAATGCCCATGCGGACACTGTGAAGTCTGTGAGAGTGTG 1140
QY 1141 TTTGGGAGCTCCGAGCGCAACCATGATGTCCCGACCTCATCAGATCGACCTGCGCC 1200
DB 1141 TTTGGGAGCTCCGAGCGCAACCATGATGTCCCGACCTCATCAGATCGACCTGCGCC 1200
QY 1201 AACCCCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 AACCCCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 GACTGCTCTGAGCAACCCAGAGCCCATCTCCCTGCGCGGAGTCTGCGCGGCGCC 1320
DB 1261 GACTGCTCTGAGCAACCCAGAGCCCATCTCCCTGCGCGGAGTCTGCGCGGCGCC 1320
QY 1321 AGCTACACCTGAGCGGAGCTGCGAGCTGGCTTTTGGCGTGGCTCCCAAGCCCTGCTCT 1380
DB 1321 AGCTACACCTGAGCGGAGCTGCGAGCTGGCTTTTGGCGTGGCTCCCAAGCCCTGCTCT 1380
QY 1381 TACATGCACTGCTGCAACCAAGCTGTGGTGACCGGGAAGGCCAAGGAGACAGATGCTGTC 1440
DB 1381 TACATGCACTGCTGCAACCAAGCTGTGGTGACCGGGAAGGCCAAGGAGACAGATGCTGTC 1440
QY 1441 CAGACCGCCACTTCCCTCGGCGGATGGCAACAGCTGTGGCGAGGCGCAAGCTCTGCTC 1500
DB 1441 CAGACCGCCACTTCCCTCGGCGGATGGCAACAGCTGTGGCGAGGCGCAAGCTCTGCTC 1500
QY 1501 AAAGGGCGCTGCGTGGAGAGACACAACTCAACAAGCACAGGCTGATGCTTCTGGGCG 1560
DB 1501 AAAGGGCGCTGCGTGGAGAGACACAACTCAACAAGCACAGGCTGATGCTTCTGGGCG 1560
QY 1561 AAATGGGATCCCTATGCGCCCTGCTCGCGACATGTGTGGTGGGCGCTGCAAGCTGGCGAG 1620
DB 1561 AAATGGGATCCCTATGCGCCCTGCTCGCGACATGTGTGGTGGGCGCTGCAAGCTGGCGAG 1620
QY 1621 AGGCACTGCAACCAACCCCTGCGCAACGGGCGCAAGTACTGCGAGGAGTGAAGGCTG 1680
DB 1621 AGGCACTGCAACCAACCCCTGCGCAACGGGCGCAAGTACTGCGAGGAGTGAAGGCTG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCCTCCGGAAGAGCTTCCGG 1740
DB 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCCTCCGGAAGAGCTTCCGG 1740

1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAACACCGGCTCACTCTGGCC 1800
1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAACACCGGCTCACTCTGGCC 1800
1801 GTGGCATGGTGCCCAAGTACTCGGCGTGTCTCCCGGGACAGTGTCAAGCTCACTCTGC 1860
1801 GTGGCATGGTGCCCAAGTACTCGGCGTGTCTCCCGGGACAGTGTCAAGCTCACTCTGC 1860
1861 CGAGCCAAATGSCACTGGCTACTTCTATGTGCTGTGSCACCCCAAGTGTGTGGACGCGCTG 1920
1861 CGAGCCAAATGSCACTGGCTACTTCTATGTGCTGTGSCACCCCAAGTGTGTGGACGCGCTG 1920
1921 TGCTCTCTGACTCCACTCGCTGTGTGTGTCGAAGGCAAGTGCATCAAGGCTGCTGTGAT 1980
1921 TGCTCTCTGACTCCACTCGCTGTGTGTGTCGAAGGCAAGTGCATCAAGGCTGCTGTGAT 1980
1981 GGGAACTGGGCTCCAAGAGAGATTCGACAAGTGTGGGCTGTGTGGGAGACAATAAG 2040
1981 GGGAACTGGGCTCCAAGAGAGATTCGACAAGTGTGGGCTGTGTGGGAGACAATAAG 2040
2041 AGCTGCAAGAAGTGTGACTGGACTTTCACCAAGCCCATGCTACATTTTCGTGGTG 2100
2041 AGCTGCAAGAAGTGTGACTGGACTTTCACCAAGCCCATGCTACATTTTCGTGGTG 2100
2101 GCCATCCCGGAGCGGCTCAAGATCGATCGCGCAGCGCGGTTACAAAGGCTGATC 2160
2101 GCCATCCCGGAGCGGCTCAAGATCGATCGCGCAGCGCGGTTACAAAGGCTGATC 2160
2161 GGGATGACAACTACCTGGCTCTGAAGACAGCAAGGCAAGTACCTGTCAACGGGCAAT 2220
2161 GGGATGACAACTACCTGGCTCTGAAGACAGCAAGGCAAGTACCTGTCAACGGGCAAT 2220
2221 TTCTGTGTGTGCGGCTGAGGCGGACCTGCTGTGTAAGGGCAAGTACCTGTGCGGTACAGC 2280
2221 TTCTGTGTGTGCGGCTGAGGCGGACCTGCTGTGTAAGGGCAAGTACCTGTGCGGTACAGC 2280
2281 GGCACGGCACAGGCTGAGGCGGCTGAGGCTTCCGGGCCCATCTCGAGCGGCTGAGC 2340
2281 GGCACGGCACAGGCTGAGGCGGCTGAGGCTTCCGGGCCCATCTCGAGCGGCTGAGC 2340
2341 GTGGAGTCTCTCCGTGGGGAAGTACACACCGCCCGGCTCGCTACTCTCTATCTG 2400
2341 GTGGAGTCTCTCCGTGGGGAAGTACACACCGCCCGGCTCGCTACTCTCTATCTG 2400
2401 CCCAAGAGCCTCGGAGGACAAGTCTCTATCCCAAGACCCCGGGGACCCCTCTGTC 2460
2401 CCCAAGAGCCTCGGAGGACAAGTCTCTATCCCAAGACCCCGGGGACCCCTCTGTC 2460
2461 TTGCACAACAGCTCTCAGCCTCTCCAAAGTGTGAGCGGACGCGGACGAGGCCCCCT 2520
2461 TTGCACAACAGCTCTCAGCCTCTCCAAAGTGTGAGCGGACGCGGACGAGGCCCCCT 2520
2521 GCACGTGGGTGGCTGGCAGCTGGGGCGGTGTCTCCGAGGTGCGGCGAGTGGCCTGCGAG 2580
2521 GCACGTGGGTGGCTGGCAGCTGGGGCGGTGTCTCCGAGGTGCGGCGAGTGGCCTGCGAG 2580
2581 AAGGGCGGTGACATGTGCGGGCTCGGGGCGGACGCGGACGCTCCCTGCTGTGATGA 2640
2581 AAGGGCGGTGACATGTGCGGGCTCGGGGCGGACGCGGACGCTCCCTGCTGTGATGA 2640
2641 GCCCATCGGCGGCTGGAGACACAAGCTCGGGGAGCCCTGCCCCACCTGGGAGCTCAGC 2700
2641 GCCCATCGGCGGCTGGAGACACAAGCTCGGGGAGCCCTGCCCCACCTGGGAGCTCAGC 2700
2701 GCTGTGTACCTGTCTCAAGAGCTCGGCGCGGGGATTTACAGAGCGCTCACTCAAGTGT 2760
2701 GCTGTGTACCTGTCTCAAGAGCTCGGCGCGGGGATTTACAGAGCGCTCACTCAAGTGT 2760
2761 GTGCGCACCGGAGCGGCTGTGGCGCGGACGAGTGAACCTTGACCGCAAGCCCGAC 2820
2761 GTGCGCACCGGAGCGGCTGTGGCGCGGACGAGTGAACCTTGACCGCAAGCCCGAC 2820
2821 GAGCTGGACTTCTGCGCTCTGAGCGCGGTGCTGA 2853

Db 2821 GAGCTGGACTTCTGCGTCTGAGGCGGTGCTGA 2853

RESULT 4

US-60-453-135-7278
; Sequence 7278, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7278
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-7278

Query Match 99.8%; Score 2848.2; DB 11; Length 2853;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2844; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCACTCAACCTGGCTTCCCGGGCGAACCCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGCACTCAACCTGGCTTCCCGGGCGAACCCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGTAGTCTGTTCCCATCCGACTGGACCCCGGACATTAACGGCGCGCTAC 120
Db 61 CCAGAGCGGGAGTAGTCTGTTCCCATCCGACTGGACCCCGGACATTAACGGCGCGCTAC 120
QY 121 TACTGCGGGGTCCGGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCAATT 180
Db 121 TACTGCGGGGTCCGGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCAATT 180
QY 181 CAGGAGGACTTTTACCTACCTGACGCGGATGCTCAGTTCTTGGTCCCGGCTCTCC 240
Db 181 CAGGAGGACTTTTACCTACCTGACGCGGATGCTCAGTTCTTGGTCCCGGCTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGCTCTTCCAGACCTCGCAGC 300
Db 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGCTCTTCCAGACCTCGCAGC 300
QY 301 TGCTTCTATTCTGGGAGCTGAACCGCGGAGCTGCTTCGCTGCTGTGAGCCTGTC 360
Db 301 TGCTTCTATTCTGGGAGCTGAACCGCGGAGCTGCTTCGCTGCTGTGAGCCTGTC 360
QY 361 GGGGGCTTCCCGGAGCCTTTGGCTACCGAGCGCGGATGTATGTCATTAGCCCGCTGCC 420
Db 361 GGGGGCTTCCCGGAGCCTTTGGCTACCGAGCGCGGATGTATGTCATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGCGGCGAGCGCAACAGCGAGCGGCGACACCTTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGGCGGCGAGCGCAACAGCGAGCGGCGACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCGGGCGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCGGGCGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTTACGGGCGCTTACAGCGCGGCGGCGGGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTTACGGGCGCTTACAGCGCGGCGGCGGGGCTTCGGGGAG 600
QY 601 AGTGTAGCGCGCGGCGGCTGTGGCGCGGCAAGGTTTCTGTCTATCCCGCGGTACGTG 660
Db 601 AGTGTAGCGCGCGGCGGCTGTGGCGCGGCAAGGTTTCTGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGTCGGGAGGAGTCAATGTTCAAGTTCACAGCGCGGACCTTGGAAAT 720
Db 661 GAGACGCTGTGTGTCGGGAGGAGTCAATGTTCAAGTTCACAGCGCGGACCTTGGAAAT 720

Db 661 GAGACGTGGTGGTCCGGGACAGTCAATGGTCAAGTTCACCGCCGGCAATCGAACAT 720
Qy 721 TATCTGTGACGTGTGGTGAACAGCGCGGAGCTCTACCGCCCATCCCAAGATCTCTCAAC 780
Db 721 TATCTGTGACGTGTGGTGAACAGCGCGGAGCTCTACCGCCCATCCCAAGATCTCTCAAC 780
Qy 781 CCCATCAACATCTGTGTGTCGAAGGTGCTGCTTTAGAGATCGTGACTCCGGGCCAAG 840
Db 781 CCCATCAACATCTGTGTGTCGAAGGTGCTGCTTTAGAGATCGTGACTCCGGGCCAAG 840
Qy 841 GTACACGGCAATCGCGGCTGACGCTCGCAACTCTCTGCTGCGCAGAGAGCTGGAAC 900
Db 841 GTACACGGCAATCGCGGCTGACGCTCGCAACTCTCTGCTGCGCAGAGAGCTGGAAC 900
Qy 901 AAAGTGTGACAAAGCACCAGTACTGGACACTGCCATCTCTTACACAGGAGGAC 960
Db 901 AAAGTGTGACAAAGCACCAGTACTGGACACTGCCATCTCTTACACAGGAGGAC 960
Qy 961 CTGTGTGGAGCCACCACTGTGACACCTGGGCATGGCTGATGGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGACACCTGGGCATGGCTGATGGGTACCATGTGTGAC 1020
Qy 1021 CCCAAGAGAGCTGCTGTGATTTGAGGACCATGGCTTTCATCAGGCTTCACCACTGCC 1080
Db 1021 CCCAAGAGAGCTGCTGTGATTTGAGGACCATGGCTTTCATCAGGCTTCACCACTGCC 1080
Qy 1081 CACGAGCTGGGCCACCTGTTCACATGCCCATGACATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CACGAGCTGGGCCACCTGTTCACATGCCCATGACATGTGAAGTCTGTGAGGAGGTG 1140
Qy 1141 TTTGGAGGCTCCGAGCAACACATGATGTCCCGGACCTTCATCCAGATCGACCTGGC 1200
Db 1141 TTTGGAGGCTCCGAGCAACACATGATGTCCCGGACCTTCATCCAGATCGACCTGGC 1200
Qy 1201 AACCCCTGTGACGCTGAGTGTGCTGATCATACCGACTTCTGTGACAGCGGACGGT 1260
Db 1201 AACCCCTGTGACGCTGAGTGTGCTGATCATACCGACTTCTGTGACAGCGGACGGT 1260
Qy 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTGCGCGCGC 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTGCGCGCGC 1320
Qy 1321 AGCTACACCTGTGACGAGCTGCGAGCTGGCTTTGGCGTGGCTTCCAGGCTCTGCTCT 1380
Db 1321 AGCTACACCTGTGACGAGCTGCGAGCTGGCTTTGGCGTGGCTTCCAGGCTCTGCTCT 1380
Qy 1381 TACATGCACTACTGCAACAGCTGTGGTGCACCGGGAAGGCAAGGCAAGATGGTGTGC 1440
Db 1381 TACATGCACTACTGCAACAGCTGTGGTGCACCGGGAAGGCAAGGCAAGATGGTGTGC 1440
Qy 1441 CAGACCCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGCGAGGGCTAAGCTCTGCC 1500
Db 1441 CAGACCCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGCGAGGGCTAAGCTCTGCC 1500
Qy 1501 AAAGGGGCTCGGTGGAGAGACACACCTCAACAGAGAGAGTGGATGTTCTGGGCG 1560
Db 1501 AAAGGGGCTCGGTGGAGAGACACACCTCAACAGAGAGAGTGGATGTTCTGGGCG 1560
Qy 1561 AAATGGGATCCCTATGGCCCTGCTCGCGCACATGTGTGCGGCGTGCAGCTGGCCAG 1620
Db 1561 AAATGGGATCCCTATGGCCCTGCTCGCGCACATGTGTGCGGCGTGCAGCTGGCCAG 1620
Qy 1621 AGGAGTGCACCAACCCAGCTGCGCAACGGGGGCAAGTACGCGAGGAGTGGAGTGT 1680
Db 1621 AGGAGTGCACCAACCCAGCTGCGCAACGGGGGCAAGTACGCGAGGAGTGGAGTGT 1680
Qy 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGGCTCCGGAAAGAGTTCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGGCTCCGGAAAGAGTTCGG 1740
Qy 1741 GAGGAGCAGTGTGAGGCTTTCAAGGCTTACAACGACACACACCGGCTCACTCTCGCC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCAAGGCTTACAACGACACACACCGGCTCACTCTCGCC 1800

Qy 1801 GTGGCATGGTGCCCAAGTACTCCGGCGTGTCTCCCGGACAAAGTGAAGCTCATCTGC 1860
Db 1801 GTGGCATGGTGCCCAAGTACTCCGGCGTGTCTCCCGGACAAAGTGAAGCTCATCTGC 1860
Qy 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGACCCCAAGGTGTGACGCGACGTG 1920
Db 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGACCCCAAGGTGTGACGCGACGTG 1920
Qy 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Qy 1981 GGGAACTGGGCTCCCAAGAGAGATTCCAGAGTGTGGGTGTGGGGAGACAAATAG 2040
Db 1981 GGGAACTGGGCTCCCAAGAGAGATTCCAGAGTGTGGGTGTGGGGAGACAAATAG 2040
Qy 2041 AGCTGCAAGAGGTGACTGGACTCTTCAAGGCCATGCATGGCTTACAAATTTCTGGTG 2100
Db 2041 AGCTGCAAGAGGTGACTGGACTCTTCAAGGCCATGCATGGCTTACAAATTTCTGGTG 2100
Qy 2101 GCCATCCCCGAGCGCTCAAGCATTCGCGCAGCGCGTTACAAAGGCGCTGATC 2160
Db 2101 GCCATCCCCGAGCGCTCAAGCATTCGCGCAGCGCGTTACAAAGGCGCTGATC 2160
Qy 2161 GGGATGACAACTACCTGGCTGTGAAGACAGCCAAAGCAAGTACCTGCTCAACGGGCA 2220
Db 2161 GGGATGACAACTACCTGGCTGTGAAGACAGCCAAAGCAAGTACCTGCTCAACGGGCA 2220
Qy 2221 TTCTGTGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGTGCGGTACAGC 2280
Db 2221 TTCTGTGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGTGCGGTACAGC 2280
Qy 2281 GCGAGGGCACAGCGGTGGAGAGCTTCGCGGCCCATCTCTGGAGCGCGTGAAC 2340
Db 2281 GCGAGGGCACAGCGGTGGAGAGCTTCGCGGCCCATCTCTGGAGCGCGTGAAC 2340
Qy 2341 GTGAGGTCTCTCTCGTGGGGAAGATGACACCGCCCGGCTCCGCTACTCTTCTATCTG 2400
Db 2341 GTGAGGTCTCTCTCGTGGGGAAGATGACACCGCCCGGCTCCGCTACTCTTCTATCTG 2400
Qy 2401 CCCAAAGAGCTCGGGGAGACAAGTCTCTCATCCAAAGGACCCCGGGGACCTCTGTG 2460
Db 2401 CCCAAAGAGCTCGGGGAGACAAGTCTCTCATCCAAAGGACCCCGGGGACCTCTGTG 2460
Qy 2461 TTGCACACAGCTCTCTCAGCTCTCCAAAGGTGGAGCGGACCGGACAGCCCGCT 2520
Db 2461 TTGCACACAGCTCTCTCAGCTCTCCAAAGGTGGAGCGGACCGGACAGCCCGCT 2520
Qy 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGGTCTCCGCGAGCTGGCGAGTGGCGTGCAG 2580
Db 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGGTCTCCGCGAGCTGGCGAGTGGCGTGCAG 2580
Qy 2581 AAGCGGGCGGTGGAATGTGCGGCTCCGCGGACAGCGCTCCCTGCTGTGTGATGCA 2640
Db 2581 AAGCGGGCGGTGGAATGTGCGGCTCCGCGGACAGCGCTCCCTGCTGTGTGATGCA 2640
Qy 2641 GCGCATCGGCGGTGGAGACACAAGCTCGGGGAGCGCTGCCCGACCTGGGAGCTCAGC 2700
Db 2641 GCGCATCGGCGGTGGAGACACAAGCTCGGGGAGCGCTGCCCGACCTGGGAGCTCAGC 2700
Qy 2701 GCTGTGTCACCTCTCTCCAAGAGCTGCGGCGGGGATTTCAAGGCGCTCACTCAAGTGT 2760
Db 2701 GCTGTGTCACCTCTCTCCAAGAGCTGCGGCGGGGATTTCAAGGCGCTCACTCAAGTGT 2760
Qy 2761 GTGGGCCAGGAGCGCGGTGTGGCCCGGGACAGTGCACACTTGCACCGACCGCCAG 2820
Db 2761 GTGGGCCAGGAGCGCGGTGTGGCCCGGGACAGTGCACACTTGCACCGACCGCCAG 2820
Qy 2821 GAGCTGGAGTCTCTCGCTCTCAGGCGCGGTGTGA 2853
Db 2821 GAGCTGGAGTCTCTCGCTCTCAGGCGCGGTGTGA 2853


```
RESULT 5
US-60-453-050-7278
; Sequence 7278, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7278
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-7278

Query Match      99.8%; Score 2848.2; DB 11; Length 2853;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2844; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ATGCTTCTGCTGGGATCCTAACCTGGCTTCGCCGGGGAACCGCTGGAGGCTCTCAG 60
Db 1 ATGCTTCTGCTGGGATCCTAACCTGGCTTCGCCGGGGAACCGCTGGAGGCTCTCAG 60

Oy 61 CCAGAGCGGAGGTAGTGTTCATCCGACTGGACCCGAGACATTAAAGCGCCGGCTAC 120
Db 61 CCAGAGCGGAGGTAGTGTTCATCCGACTGGACCCGAGACATTAAAGCGCCGGCTAC 120

Oy 121 TACTGGCGGGTCCCGAGGATCCCGGGGATCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCCGAGGATCCCGGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180

Oy 181 CAGGAGGACTTTTACTACACTGACCCGGATGCTCAGTTCTTGCTCCGCTCCGCTCC 240
Db 181 CAGGAGGACTTTTACTACACTGACCCGGATGCTCAGTTCTTGCTCCGCTCCGCTCC 240

Oy 241 ACTGAGATCTGGGCTGCCCTCCAGGGGTACCGGGGTCTACAGGCTGCTGACGACG 300
Db 241 ACTGAGATCTGGGCTGCCCTCCAGGGGTACCGGGGTCTCTACAGGCTGCTGACGACG 300

Oy 301 TGCTTCTATTCTGGGAGGTGAAGCCGAGCCGAGCTCGCTTCGCTGCTGAGCCTGTC 360
Db 301 TGCTTCTATTCTGGGAGGTGAAGCCGAGCCGAGCTCGCTTCGCTGCTGAGCCTGTC 360

Oy 361 GGGGGCTCCGGGAGCTTTGGCTACGAGCGCGAGTATGCTATAGCCCGCTGCC 420
Db 361 GGGGGCTCCGGGAGCTTTGGCTACGAGCGCGAGTATGCTATAGCCCGCTGCC 420

Oy 421 AATGCTAGCGCGCGGCGGAGCGGACGAGCCAGGCGGCTACCTTCTCCAGCGCCGG 480
Db 421 AATGCTAGCGCGCGGCGGAGCGGACGAGCCAGGCGGCTACCTTCTCCAGCGCCGG 480

Oy 481 GGTGTTCCGGGCGGCTTCCGGAGACCCACCTCTCGTTCGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCCGGGCGGCTTCCGGAGACCCACCTCTCGTTCGGGGTGGCTCGGGCTGG 540

Oy 541 AACCCCGCATCTACGGGCTGAGCCCTTACAAGCCGCGCGGCGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTACGGGCTGAGCCCTTACAAGCCGCGCGGCGGCTTCGGGGAG 600

Oy 601 AGTCGTAGCGCGGAGCTCTGGGCGGCGGAGCGGCTTCTATCCCGGCTAGCTG 660
Db 601 AGTCGTAGCGCGGAGCTCTGGGCGGCGGAGCGGCTTCTATCCCGGCTAGCTG 660

Oy 661 GAGACGCTGGTGTGCGGAGGAGTCAATGTTCAAGTTCACGGCGGCGGAGCTGGAACAT 720
Db 661 GAGACGCTGGTGTGCGGAGGAGTCAATGTTCAAGTTCACGGCGGCGGAGCTGGAACAT 720

Oy 721 TATCTGCTGACGCTGCTGGCAACGGCGCGGACTCTACCGGCTATCCAGCATCTCAAC 780
Db 721 TATCTGCTGACGCTGCTGGCAACGGCGCGGACTCTACCGGCTATCCAGCATCTCAAC 780
```

```
Db 721 TATCTGCTGACGCTGCTGGCAACGGCGCGGACTCTACCGGCTATCCAGCATCTCAAC 780
Oy 781 CCATCAACATCGTTGTGCTCAAGGTGCTGCTTCTTAGAGATCGTGACTCCGGGCGCAAG 840
Db 781 CCATCAACATCGTTGTGCTCAAGGTGCTGCTTCTTAGAGATCGTGACTCCGGGCGCAAG 840
Oy 841 GTCCACGGCAATGGCGGCTGAGCTGCGCAACTTCTGTGCTGCGCAAGCAAGCTGAAC 900
Db 841 GTCCACGGCAATGGCGGCTGAGCTGCGCAACTTCTGTGCTGCGCAAGCAAGCTGAAC 900
Oy 901 AAAGTAGAGTGAAGACACCCCGAGTACTGGGACACTGCGCATCTTCCACAGGAGGAC 960
Db 901 AAAGTAGAGTGAAGACACCCCGAGTACTGGGACACTGCGCATCTTCCACAGGAGGAC 960
Oy 961 CTGTGTGGAGCCACACCTGTGTGACACCCCTGGCATGGCTGTGTGGTGTGTGTGAC 1020
Db 961 CTGTGTGGAGCCACACCTGTGTGACACCCCTGGCATGGCTGTGTGGTGTGTGTGAC 1020
Oy 1021 CCCAAGAGAAGTCTCTGTGCTATTGAGGAGTGGGCTTCCATCAGCCTTCCACACTGCC 1080
Db 1021 CCCAAGAGAAGTCTCTGTGCTATTGAGGAGTGGGCTTCCATCAGCCTTCCACACTGCC 1080
Oy 1081 CAGGAGCTGGGCGGCTGTTTCAACATGCCCATGCAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CAGGAGCTGGGCGGCTGTTTCAACATGCCCATGCAATGTGAAGTCTGTGAGGAGGTG 1140
Oy 1141 TTTGGGAAGCTCCGAGCCAAACCATGATGTCCCGACCCCTATCCAGATCGACCGTGC 1200
Db 1141 TTTGGGAAGCTCCGAGCCAAACCATGATGTCCCGACCCCTATCCAGATCGACCGTGC 1200
Oy 1201 AACCCCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 AACCCCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Oy 1261 GACTGCTGCTGAGGACACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1320
Db 1261 GACTGCTGCTGAGGACACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1320
Oy 1321 AGTACACCTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1380
Db 1321 AGTACACCTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1380
Oy 1381 TACATGCACTGCTGACCAAGCTGTGTGTCACCGGGAAGGCAAGGGACAGATGCTGTGC 1440
Db 1381 TACATGCACTGCTGACCAAGCTGTGTGTCACCGGGAAGGCAAGGGACAGATGCTGTGC 1440
Oy 1441 CAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 CAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Oy 1501 AAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1501 AAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Oy 1561 AAATGGGATCCCTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db 1561 AAATGGGATCCCTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Oy 1621 AGGAGTGCACCAACCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db 1621 AGGAGTGCACCAACCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Oy 1681 AAATACCGATCTGCAATCTGGAGGCTTCCCGGAGTCCCGGAGTCCCGAAGAGCTTCCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGGCTTCCCGGAGTCCCGGAGTCCCGAAGAGCTTCCGG 1740
Oy 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1800
Oy 1801 GTGGGATGGTGGCAAGTACTCCGGCTGTCTCCCGGAGACAGTGCAGCTCATCTGCTGCT 1860
Db 1801 GTGGGATGGTGGCAAGTACTCCGGCTGTCTCCCGGAGACAGTGCAGCTCATCTGCTGCT 1860
```

Db 1801 GTGGCATGGTGGCCCAAGTACTCGGGGTGTCTCCCGGGCAAGTGCACAGCTCATCTGC 1860
QY 1861 CGAGCAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGTGGAGCGGACGCTG 1920
Db 1861 CGAGCAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGTGGAGCGGACGCTG 1920
QY 1921 TGCTTCTCTGACTCCAGCTCGTGTGTGTCCAAAGGCAAGTGCATCAAGCTGGCTGTGAT 1980
Db 1921 TGCTTCTCTGACTCCAGCTCGTGTGTGTCCAAAGGCAAGTGCATCAAGCTGGCTGTGAT 1980
QY 1981 GGGAACTTGGGCTCCAAAGAGAGATTGCAAGTGTGGGTGTGTGGGGAGAGCAATAAG 2040
Db 1981 GGGAACTTGGGCTCCAAAGAGAGATTGCAAGTGTGGGTGTGTGGGGAGAGCAATAAG 2040
QY 2041 AGCTGCAGAGAGTGTGACTGTGACTTCTCACCAGCCCATGCATGGCTACAAATTTCGTGGTG 2100
Db 2041 AGCTGCAGAGAGTGTGACTGTGACTTCTCACCAGCCCATGCATGGCTACAAATTTCGTGGTG 2100
QY 2101 GCCATCCCGAGGCGCTCAAGCATCGACATCCGCCAGCGGGTTACAAGGGCTGATC 2160
Db 2101 GCCATCCCGAGGCGCTCAAGCATCGACATCCGCCAGCGGGTTACAAGGGCTGATC 2160
QY 2161 GGGATGACAACTACCTGGCTCTCAAGACAGCCAGCAAGTACCTGCTCAAGGGCAT 2220
Db 2161 GGGATGACAACTACCTGGCTCTCAAGACAGCCAGCAAGTACCTGCTCAAGGGCAT 2220
QY 2221 TTCGTGGTGTCCGGGTGGAGCGGACCTGGTGGTGAAGGAGTCTGCTGGGTACAGC 2280
Db 2221 TTCGTGGTGTCCGGGTGGAGCGGACCTGGTGGTGAAGGAGTCTGCTGGGTACAGC 2280
QY 2281 GGCAGCGCACAGGGTGGAGAGCTGCAAGCTTCCCGGCCCATCTGGAGCGCGTGACC 2340
Db 2281 GGCAGCGCACAGGGTGGAGAGCTGCAAGCTTCCCGGCCCATCTGGAGCGCGTGACC 2340
QY 2341 GTGGAGGTCTCTCCGTGGGGAAGTACACACCGCCCGGTCCGCTACTCTCTATCTG 2400
Db 2341 GTGGAGGTCTCTCCGTGGGGAAGTACACACCGCCCGGTCCGCTACTCTCTATCTG 2400
QY 2401 CCCAAGAGCTCGGGAGGAGCAAGTCTCTCATCCCAAGCCCGGGAGCCCTCTGTC 2460
Db 2401 CCCAAGAGCTCGGGAGGAGCAAGTCTCTCATCCCAAGCCCGGGAGCCCTCTGTC 2460
QY 2461 TTGCAACAAGCTCTCAGCTCTCAACCAAGTGGAGAGCGGAGCAAGGCGGCGCT 2520
Db 2461 TTGCAACAAGCTCTCAGCTCTCAACCAAGTGGAGAGCGGAGCAAGGCGGCGCT 2520
QY 2521 GCACGCTGGTGGTGGCAGCTGGGGCCGTGCTCCGCACTGGCGGAGTGGCCTGCAG 2580
Db 2521 GCACGCTGGTGGTGGCAGCTGGGGCCGTGCTCCGCACTGGCGGAGTGGCCTGCAG 2580
QY 2581 AAGCGGGGTGGAGTGTGGGGTCCCGGGAGGAGCAAGTCCCTGCTGTGATGCA 2640
Db 2581 AAGCGGGGTGGAGTGTGGGGTCCCGGGAGGAGCAAGTCCCTGCTGTGATGCA 2640
QY 2641 GCCATCGGCGCGTGGAGACAAAGCTGGGGAGCCCTGCGCCACCTGGGAGCTCAGC 2700
Db 2641 GCCATCGGCGCGTGGAGACAAAGCTGGGGAGCCCTGCGCCACCTGGGAGCTCAGC 2700
QY 2701 GCTGTGACCTGCTCCAGAGCTGCGCGGGGATTTTCAGAGGCTCACTCAAGTGT 2760
Db 2701 GCTGTGACCTGCTCCAGAGCTGCGCGGGGATTTTCAGAGGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGAGGAGCGGCTGCTGGCCCGGGAGCAAGTGCACGCAAGGCCUAG 2820
Db 2761 GTGGGCGAGGAGCGGCTGCTGGCCCGGGAGCAAGTGCACGCAAGGCCUAG 2820
QY 2821 GAGCTGGACTTCTGGTCTCTGAGCGCGTGTGTA 2853
Db 2821 GAGCTGGACTTCTGGTCTCTGAGCGCGTGTGTA 2853

; Sequence 32, Application US/10311035
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Aneena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEUANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,9
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7473089CB1
; US-10-311-035-32

Query Match 98.8%; Score 2817.6; DB 9; Length 2930;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2846; Conservative 0; Mismatches 4; Indels 6; Gaps 2;
QY 1 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCCCGGGGCAACGCTGGAGGCTGTGAG 60
Db 75 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCCCGGGGCAACGCTGGAGGCTGTGAG 134
QY 61 CCAGAGCGGGAGGTAGTGGTTCCTCCACTGGAGCCCGGAGCATTAACGGCGGAGTAC 120
Db 135 CCAGAGCGGGAGGTAGTGGTTCCTCCACTGGAGCCCGGAGCATTAACGGCGGAGTAC 194
QY 121 TACTGGCGGGTCCCGGAGGACTCCGGGGATCAGGAGCTCATTTTCAGATCACAGCATTT 180
Db 195 TACTGGCGGGTCCCGGAGGACTCCGGGGATCAGGAGCTCATTTTCAGATCACAGCATTT 254
QY 181 CAGGAGGACTTTTACCTACCTAGCCCGGATGCTAGTTCTTGGCTCCCGCTTCTCC 240
Db 255 CAGGAGGACTTTTACCTACCTAGCCCGGATGCTAGTTCTTGGCTCCCGCTTCTCC 314
QY 241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCAGGGGCTTTCAGACCTTCGAGCG 300
Db 315 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCAGGGGCTTTCAGACCTTCGAGCG 374
QY 301 TCGTCTTATTCTGGGAGGTGAAGCGCGGAGCTCGTTCGCTGCTGTGAGCTGTGTC 360
Db 375 TCGTCTTATTCTGGGAGGTGAAGCGCGGAGCTCGTTCGCTGCTGTGAGCTGTGTC 434

QY 361 GGGGGCTCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCTATTAAGCCGCTGCC 420
DB 435 GGGGGCTCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCTATTAAGCCGCTGCC 494
QY 421 AATGCTAGCGCGCGCGCGCGCGAGCGAAGCAGCGAGGCGCGACACCTTCTCAGAGCGCG 480
DB 495 AATGCTAGCGCGCGCGCGCGCGAGCGAAGCAGCGAGGCGCGACACCTTCTCAGAGCGCG 554
QY 481 GGTGTTCCGGCGCGCGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 540
DB 555 GGTGTTCCGGCGCGCGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 614
QY 541 AACCCGCGCATCTACGGCGCGCTGGACCCCTTACAGCGCGCGCGCGCGCTGGGGAG 600
DB 615 AACCCGCGCATCTACGGCGCGCTGGACCCCTTACAGCGCGCGCGCGCGCTGGGGAG 674
QY 601 AGTGTAGCGCGCGCGCTGGCGCGCGCAAGCGTTTGTGTATCCCGCGGTACGCTG 660
DB 675 AGTGTAGCGCGCGCGCTGGCGCGCGCAAGCGTTTGTGTATCCCGCGGTACGCTG 734
QY 661 GAGACGCTGGTGGCGCGCGAGTCAATGTCAGGTTCCAGCGCGCGAGCTGGAAAT 720
DB 735 GAGACGCTGGTGGCGCGCGAGTCAATGTCAGGTTCCAGCGCGCGAGCTGGAAAT 794
QY 721 TATCTGCTACGCTGCTGGCAACGCGCGCGAGCTTACCGCATCCAGACATCTCAAC 780
DB 795 TATCTGCTACGCTGCTGGCAACGCGCGCGAGCTTACCGCATCCAGACATCTCAAC 854
QY 781 CCCATCAACATCGTTGGTCAAGTGCTGCTTTAGAGATCGTGACTCCCGGCGCCAG 840
DB 855 CCCATCAACATCGTTGGTCAAGTGCTGCTTTAGAGATCGTGACTCCCGGCGCCAG 914
QY 841 GTCACCGCAATGCGGCGCTGACGCTGCGCACTTCTGCTGCGCGAGAGAGCTGAAC 900
DB 915 GTCACCGCAATGCGGCGCTGACGCTGCGCACTTCTGCTGCGCGAGAGAGCTGAAC 974
QY 901 AAGTGAAGTACAGACCGCGAGTACTGGGACACTGCGCATCTTTCACAGGCGAGAC 960
DB 975 AAGTGAAGTACAGACCGCGAGTACTGGGACACTGCGCATCTTTCACAGGCGAGAC 1034
QY 961 CTGTGTGGAGCCACCATCTGTACACCGCTGGGATGCTGATGGGTTACCATGTGTGAC 1020
DB 1035 CTGTGTGGAGCCACCATCTGTACACCGCTGGGATGCTGATGGGTTACCATGTGTGAC 1094
QY 1021 CCCAAGAGAGCTGCTGTCTATTTAGAGAGATGGGCTTCCATCAGCCTTACCACTGCC 1080
DB 1095 CCCAAGAGAGCTGCTGTCTATTTAGAGAGATGGGCTTCCATCAGCCTTACCACTGCC 1154
QY 1081 CAGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAGTCTCTCAGAGAGTG 1140
DB 1155 CAGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAGTCTCTCAGAGAGTG 1214
QY 1141 TTTGGGAAGCTCCGAGCAACACATGATGTGCCGACCTCATCCAGATCGACCTGACC 1200
DB 1215 TTTGGGAAGCTCCGAGCAACACATGATGTGCCGACCTCATCCAGATCGACCTGACC 1274
QY 1201 AACCCCTGTGAGCTGAGTGTGCTGCTATCATCAGGACTTCTGAGACGCGGCGACGCT 1260
DB 1275 AACCCCTGTGAGCTGAGTGTGCTGCTATCATCAGGACTTCTGAGACGCGGCGACGCT 1334
QY 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCCGCGCGAGGATCTGCGCGGCGCC 1320
DB 1335 GACTGCTCTCTGAGCAACCCAGCAAGCCATCTCCGCGCGAGGATCTGCGCGGCGCC 1394
QY 1321 AGCTACACCTGAGCGACGCTGGAGTGGCTTTTGGCGTGGGCTTCCAGGCGCTGCTCT 1380
DB 1395 AGCTACACCTGAGCGACGCTGGAGTGGCTTTTGGCGTGGGCTTCCAGGCGCTGCTCT 1454
QY 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGTGTGC 1440
DB 1455 TACATGAGTACTGACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGTGTGC 1514

QY 1441 CAGACCGCGCACTTCCCTGGCGCATGGCACACAGCTGTGGCGAGGCAAGCTTGTGCTC 1500
DB 1515 CAGACCGCGCACTTCCCTGGCGCATGGCACACAGCTGTGGCGAGGCAAGCTTGTGCTC 1574
QY 1501 AAAGGGCGCTGGTGGAGAGACACACCTTCAACAGCACAGGTTGGATGTTCTTGGGCC 1560
DB 1575 AAAGGGCGCTGGTGGAGAGACACACCTTCAACAGCACAGGTTGGATGTTCTTGGGCC 1634
QY 1561 AAATGGGATCCCTATGGCCCTGCTCGGCGACATGTGTGGGGCGCTGACGCTGGCCAG 1620
DB 1635 AAATGGGATCCCTATGGCCCTGCTCGGCGACATGTGTGGGGCGCTGACGCTGGCCAG 1694
QY 1621 AGGCAAGTGCACAAACCCCTGCTGCAACAGGGGCAAGTACTCGAGGAGTGAAGGTG 1680
DB 1695 AGGCAAGTGCACAAACCCCTGCTGCAACAGGGGCAAGTACTCGAGGAGTGAAGGTG 1754
QY 1681 AAATACCGATCTGCAATCTGGAGCGCTGCGCCAGCTCAGCTCCGAAAGAGCTTCCGG 1740
DB 1755 AAATACCGATCTGCAATCTGGAGCGCTGCGCCAGCTCAGCTCCGAAAGAGCTTCCGG 1814
QY 1741 GAGGAGCAGTGTGAGGCTTTCACAGGCTTACACACAGCACCAACCGGCTCACTCTGCC 1800
DB 1815 GAGGAGCAGTGTGAGGCTTTCACAGGCTTACACACAGCACCAACCGGCTCACTCTGCC 1874
QY 1801 GTGGCATGGGTCGCCAGTACTCCGCGCTGCTCCCGGAGCAAGTGCACAGCTCATCTGC 1860
DB 1875 GTGGCATGGGTCGCCAGTACTCCGCGCTGCTCCCGGAGCAAGTGCACAGCTCATCTGC 1934
QY 1861 CGAGCAATGGCACTGGCTACTTCTATGTGTGCGCACCCAA ---GGTGTGGAGCGGACG 1917
DB 1935 CGAGCAATGGCACTGGCTACTTCTATGTGTGCGCACCCAAAGGTTGGTGGAGCGGACG 1994
QY 1918 CTGTGCTCTCTGACTCCAGCTCCGCTGCTGTGCAAGGCAAGTGCATCAAGGCTGGCTGT 1977
DB 1995 CTGTGCTCTCTGACTCCAGCTCCGCTGCTGTGCAAGGCAAGTGCATCAAGGCTGGCTGT 2054
QY 1978 GATGGGACCTGGGCTCCAGAGAGATTCGACAGTGTGGGCTGTGGGGGAGCAAT 2037
DB 2055 GATGGGACCTGGGCTCCAGAGAGATTCGACAGTGTGGGCTGTGGGGGAGCAAT 2114
QY 2038 AAGAGCTGCAAGAGGTGACTGGACTCTTCAACAGCCCATGCTGCTACAAATTCGCTG 2097
DB 2115 AAGAGCTGCAAGAGGTGACTGGACTCTTCAACAGCCCATGCTGCTACAAATTCGCTG 2174
QY 2098 GTGGCATCCCGCAGCGCTCAAGCATCGACATCCCGCAGCGGTTTCAAGAGGCTG 2157
DB 2175 GTGGCATCCCGCAGCGCTCAAGCATCGACATCCCGCAGCGGTTTCAAGAGGCTG 2234
QY 2158 ATCGGGATGACAACTACCTGGCTCTGAAGAAGCACGCCAAGCAAGTACTGCTCAAGGG 2217
DB 2235 ATCGGGATGACAACTACCTGGCTCTGAAGAAGCACGCCAAGCAAGTACTGCTCAAGGG 2294
QY 2218 CATTTCTGTGTGCGCGGTGGAGCGGACCTGTGTGTAAGGGCAGTCTGCTGCGGTAC 2277
DB 2295 CATTTCTGTGTGCGCGGTGGAGCGGACCTGTGTGTAAGGGCAGTCTGCTGCGGTAC 2354
QY 2278 AGCGGACGGGACACAGGCTGGAGAGCTGACGCTTCCCGGCCCATCTCTGGAGCGCTG 2337
DB 2355 AGCGGACGGGACACAGGCTGGAGAGCTGACGCTTCCCGGCCCATCTCTGGAGCGCTG 2414
QY 2338 ACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2397
DB 2415 ACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2474
QY 2398 CTGCCAAAGAGCTCGGGAGGACAACTCTCTCATCCAAAGAGACCCCG ---GACCC 2454
DB 2475 CTGCCAAAGAGCTCGGGAGGACAACTCTCTCATCCCGCCACCCCGGGAGGACCC 2534
QY 2455 TCTGTCTTGCACAAAGCGCTCTCAACAGGTTGGAGAGCGCGGAGGACAGG 2514
DB 2535 TCTGTCTTGCACAAAGCGCTCTCAACAGGTTGGAGAGCGCGGAGGACAGG 2594
QY 2515 CCCCCTGCACGCTGGGTGGCTGGCAGCTGGGGCGGCTGCTCCGGAGCTGCGGCGAGTGGC 2574

Db 1501 AAGGAGCCTGGGTGAGAGACACAAACCAAGTACCGGGTGGACGGCTCTTGGGCG 1560
Qy 1561 AAATGGATCCTATGCGCCCTGCTCGGCACATGCTGGTGGGCGTGCAGTGGCCAGG 1620
Db 1561 AAGTGGAGCCCTACGTTCTGCTCCCGACCTGCGGGTGGGGGCTGCGAGCTGGCCCG 1620
Qy 1621 AGGCACTGACCAACCCACCCCTGCCAACGGGGGCAAGTACTGCGAGGAGTGAAGGTTG 1680
Db 1621 AGGCACTGACCAACCCCTACCCCTGCCAACGGTGAAGTACTGCGAGGAGTGAAGGTTG 1680
Qy 1681 AAATACCGATCCTGCAATCTGAGCGCTGCCCGAGCTCAACCTCCGGAAGAGCTTCCG 1740
Db 1681 AAATACCGATCCTGCAATCTGAGCGCTGCCCGAGCTCAACCTCCGGAAGAGCTTCCG 1740
Qy 1741 GAGGAGAGTGGAGCTTTTCAACGGCTTACACCCAGCAGCAACACCGGCTTCACTTCGGC 1800
Db 1741 GAGGAGAGTGGAGCTTTTCAATGGCTTACACCCAGCAGCAACACCGGCTTCACTTAGCT 1800
Qy 1801 GTGGCATGGTGGCCCAAGTACTCGGCGTGTCTCCCGGGACAAAGTCAAGCTCATCTGC 1860
Db 1801 GTGGCATGGTACCAGTACTACGGCGTGTACACAGCTGACAGTGTACAGTCACTCTGC 1860
Qy 1861 CGAGCAATGGGACTGGCTTCTATGCTGGCAGCCACCAAGTGTGGACGCGTG 1920
Db 1861 CGAGCAATGGGACTGGCTTCTATGCTAGCAGCTAAGTGTGGTGGTGGTACGCTG 1920
Qy 1921 TGCTCTCTGACTCCACCTCGCTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCGCTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Qy 1981 GGAACCTGGGCTCCAGAGAGATTCGACAAGTGTGGGGTGTGGGGGAGACATAAG 2040
Db 1981 GGAACCTGGGCTCCAGAGAGATTTGACAAATGTGGTGTGGTGGGAGACATAAG 2040
Qy 2041 AGCTGAAGAGTGTGAGTCTTCAACAGCCATGCTACATGCTACAAATTTCTGGTG 2100
Db 2041 AGCTGAAGAGTGTGAGTCTTCAACAGCCATGCTACATGCTACAAATTTCTAGTG 2100
Qy 2101 GCATCCCGAGGCGCTCAAGATTCGACATTCGCGCAGCGGCTTACAAGGGCTGATC 2160
Db 2101 GCATCCCGAGGCGCTCCAGAGTGTATGCTGAGCGTGTGCTGAGCGTGTGCTGAG 2160
Qy 2161 GGGATGACAACTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGGCA 2220
Db 2161 GGAGATGACAACTACCTGGCTTTGAAGAACAGCAAGCAAGTACCTGCTCAATGGGCA 2220
Qy 2221 TTGCTGGTGTGCGGCTGGAGCGGACCTGGTGTGAGGCGAGTCTGCTCGGCTACAG 2280
Db 2221 TTGCTGGTGTGCGGCTGGAGCGGACCTGGTGTGAGGCGAGTCTGCTACGCTATAG 2280
Qy 2281 GGCACGGGACAGCGGTGGAGAGCCTGCAAGCTTCCCGGCGCCATCTCTGGAGCGCTGAC 2340
Db 2281 GGCACCTGGCACAGTGGAGAGCCTGCAAGCTTCTCGACCCATCTCTGGAGCGCAC 2340
Qy 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACGGCGCGGCTCGCTACTCTCTATCTG 2400
Db 2341 GTGGAGTCTCTCTCCGTGGGGAAGATGACACGGCGCGGCTCTGCTATCTCTACCTG 2400
Qy 2401 CCAAGAGAGCTCGGAGGACAAAGTCTCTCATCCAGGAGACCCCGGGGACCCCTCTGTC 2460
Db 2401 CCAAGAGAGCTCGGAGGACAAAGTCTCTCATCCAGGAGACCCCTCGAGGTTCCCGGCTG 2460
Qy 2461 TTGCACAGAGCTCTCAGCTCTCCAGCCAGTGGAGAGCGGAGCGGACAGGCGCCCT 2520
Db 2461 TTGCAGAACAGTCTCTCAGCTCTTCCAGCCAGTGGAGAGCGGAGAGTACAGACCTCT 2520
Qy 2521 GCAGCTGGTGGCTGGCAGCTGGGGGCGCTGCTCCGCGAGCTGCGGAGTGGGCTGCGAG 2580
Db 2521 GCAGCTGGTGGCAGGAGCTGGGAGCTTGTCTAGTGAAGTGTGGCAGTGGCTTACAG 2580
Qy 2581 AAGCGGCGGTGAGTGTGGGGCTCCCGCGGAGCGGAGCGGCTGCTGCTGTGATGA 2640
Db 2581 AAGCGAGAGTGGAGTGTGGGGGAGTCTCCCGAGGCGGAGCAAGAGCTCTGCTTGTGATGA 2640

Qy 2641 GCCATCGGCGCTGGAGACACAAAGCCTGGGGAGCCCTGCCCCACCTGGGAGCTCAGC 2700
Db 2641 GACCATCGGCCATTGGAGAGAGCGAGCCTGTGGGAACCTGTGGAAGCTTGGAGCTCGC 2700
Qy 2701 GCTGTGTACCCCTGCTCAAGAGCTGGCGCGGGGATTTAGAGGGCTCTACTCAAGTGT 2760
Db 2701 AACTGTGTCCCTGCTTCTAAAGCTGTGGCGTGGATTTAAGAGGGCTCGCTCAAGTGT 2760
Qy 2761 GTGGGCGAGGAGCGGCTGCTGGCGCGGAGCAGTGCACCTTGCACCGCAAGCCCGAC 2820
Db 2761 GTGGGCGAGGAGCGGCTGCTGGCTGGGAGCAGTGTGACTTACGCCGCAAGCCCGAC 2820
Qy 2821 GAGCTGGACTTCTGCTCTCTGAGCGCGCTGCTGA 2853
Db 2821 GAATTAGACTTCTGCTCTTGAGACCCCTGCTGA 2853

RESULT 8

US-60-455-444-3860

; Sequence 3860, Application US/60455444

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: BEGOVICH, Ann

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; FILE REFERENCE: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001455

; CURRENT APPLICATION NUMBER: US/60/455,444

; NUMBER OF SEQ ID NOS: 50986

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3860

; LENGTH: 3933

; TYPE: DNA

; ORGANISM: Homo sapiens

US-60-455-444-3860

Query Match 23.3%; Score 664.8; DB 10; Length 3933;

Best Local Similarity 59.0%; Pred. No. 1.7e-133;

Matches 1349; Conservative 7; Mismatches 884; Indels 48; Gaps 11;

Qy 157 CTCAATTTTCAGATCAGACATTTTCAGGAGGACTTTTACCTACACCTGACGCGGATGCT 216
Db 1208 CTGCGCTCCACCTGTCCGCTTCGGCAAGGCTTCTGCTGCGCTGGCGCCGCGAGAC 1267
Qy 217 CAGTTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGCGTCCCGCTCCAGGGCTCACC 276
Db 1268 AGCTTCTTGGCGCGGCTTCAAGATGAGGCGCTTCGGGGCTCCGCGCGGCG---ACC 1324
Qy 277 GGGGGCTTTCAGACCTCGCAGCTGCTTCTATTCTGGGAGCTGAACGCGAGCGGAC 336
Db 1325 GGGGGCGAGCGGGCTGCGGGCTGCTTCTTCTCCGCGACCGCTGAATGGGAGCGCGAG 1384
Qy 337 TCGTTCGCTGCTGAGCGCTGCGGGGCTTCGCGGAGCCTTGGCTACCGAGCGGC 396
Db 1385 TCCTCGCGCGGCTGAGCTGTGCGGGCTGAGCGGCTTCTCTGCTGGACGCGGAG 1444
Qy 397 GAGTATGCTATTAGCCCGCTGCCAATGCTAGCGCGCGCGGCGGAGCGGCGCAACAGCCAG 456
Db 1445 GAGTTCAACATCCAGCGCGGCGCGGGGCTCCCTGGSTGAGCGCGCACCGCTGCGAG 1504
Qy 457 GCGCACACCTTCTCCAGCGCGGGGTGTTCCGGGCGGCGCTTCCGAGAGACCCACTCT 516
Db 1505 CGTGGGGTCCCGCGGAGCGCGCCCTCCCGAGAGACCCGAGTGGGAGTGGAGAGC 1564
Qy 517 CGTGGCGGGTG-GCCTCGGGCTGGAAACCCCGCATCTTACGGGCGCTGGACCTTACAA 575
Db 1565 GGAGAGGGTTCAGAGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624
Qy 576 GCGCGCGCGCGGGCTTCGGGGAGAGTCTGAGCGCGCGCAGG-----TCTGGCGCGCGC 630
Db 1625 GAAGAGAGGAGAGAGGCGCTAGCGAGCGCGCCACCGCCCTTGGGGGCGCAGTAGGACC 1684

QY 631 AACGTTTCTGTTCTATCCCGCGGTACGTGGAGACGCTGGTGTCCGGACAGTCAATG 590
Db 1685 AAGCGGTTTGTCTGAGGCGCTGCTGGAGACGCTGCTGGCCGATCGCTCCATG 1744
QY 691 GTCAAGTTCACGCGCGGACCTGGAAACATATCTCTGACGCTGTGTGCAACGCGGG 750
Db 1745 GCTGCCCTTTTACCGGGCGGACCTGGAGAACCATCTCTGACGTTAATGCTGTGGCAGCC 1804
QY 751 CGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCGTTGTGTCAAGTGTCTG 810
Db 1805 CGAATCTACAGCAGCCGACATCAAGAATTCATCAACCTGATGTTGGTAAAGTCTG 1864
QY 811 CTCTTTAGAGATCGTACTCCGGGCCAAGGTACCGGCAATCGCGCCCTGAGCGTCCG 870
Db 1865 ATCTAGAGATGAATAATGGGCGCCAGAGGTGTCCGCAATGGGGGCTTACACTGCGT 1924
QY 871 AACTTCTGCTCGGCAAGAAAGCTGAACAAAGTGAAGTGAAGACACCCCGAGTACTGG 930
Db 1925 AACTTCTGCACTGGCAAGCGGCTTTCAACAGCCGAGGACCCGACCAAGAGCACTAY 1984
QY 931 GACACTGCCATCTCTTACAGCAGGACCTGTGTGG----AGCCACACCTGTGACACC 987
Db 1985 GACAGCGCATCTCTGCTACAGACAGAACTTCTGTGGCAGGAGGCTGTGTGACACC 2044
QY 988 CTGGCATGGCTGATGTGGGTACCATGTGTACCCCAAGAGAGCTGCTGTGTCATTGAG 1047
Db 2045 CTGGGTGGGAGACATCGGGACCATTTGTGACCCCAACAAAGCTGCTCGTGTATCGAG 2104
QY 1048 GACGATGGGCTTCATCAGCTTCACCACTGCCACAGGTGGGCCACGTTTCAACATG 1107
Db 2105 GATGAGGGCTCCAGGCGGCCACACCTGGCCCATGAATAGGCAACGTTCTCAGCATG 2164
QY 1108 CCCATGACATGTGAAGTCTGTGAGAGGTGTTGGGAGCTCCGAGCCACACCATG 1167
Db 2165 CCCACGACGACTCCAAAGCCCTGCACACGCTCTTCGGGCCATGGGCAAGACACCATG 2224
QY 1168 ATGTCCCGACCTTCATCAGATGACGCTGCCAACCCCTGGTGCAGCTGCACTGTGCTGCC 1227
Db 2225 ATGGCAGCGCTGTTGCTCCACTGACAGAGCTGCTGCTGCTGCCCTGTCAGCGCCATG 2284
QY 1228 ATCATACCGACTTCTCTGGACAGGGGACGGTGACTGCTGCTGCTGGACCAACAGCAAG 1287
Db 2285 TATCTACAGAGCTTCTGGAGGGGGCACGAGACTGCTCTCTGGATGTCCTGCTGCTGG 2344
QY 1288 CCCATCTCCGTCGCGAGACTGCGGCG----GCCAGCTACACCTGAGCCAGCAG 1341
Db 2345 GCGCTGCCCTCCCCACAGGCTCCCGGGCTGCGATGGCCCTGTACAGCTGGACAGCAG 2404
QY 1342 TCGAGCTGGCTTTTGGGCTGGGCTCCAAAGCCCTGTCTCTTACA-----TGCAGTAC 1392
Db 2405 TCGAGGAGATCTTTGGGCGGATTTCCGCCACTGCCCCAACACCTCTGCTCAGGAGCTC 2464
QY 1393 TGCACCAAGCTGTGTG---CACCGGAGGCCAGGACAGATGGTGTGCCACACCGG 1449
Db 2465 TCGCGCCAGCTTTGTGGCCACTGATGGGCTGAGCCCTGTGCCACCAAGAAATGGC 2524
QY 1450 CACTTCCCTGGCGGATGGCAACAGCTGTGGGAGGCAAGCTGCTGCTCAAGGGGCC 1509
Db 2525 AGCTGCTGCTGGCTGAGGCAAGCGCTGCGGCTGGGCACTGCTGCTGAGAGGACG 2584
QY 1510 TG-----CGTGGAGAGACAACTCAACAGCAGAGGTGGATGGTTCCTGGGCCAAA 1563
Db 2585 TGTCTACCTGAGGAGAGTGGAGAGGCCCAAGCCGTTGGAGATGGAGGCTGGGACCG 2644
QY 1564 TGGGATCCCTATGGCCCTGCTCGGCACATGTGTGGGGGGTGCAGCTGCCAGGAGG 1623
Db 2645 TGGGGACCTTGGGGAGATGTTCTCGGACCTGTGGAGAGGAGTACAGTTTTCACACCGT 2704
QY 1624 CAGTGCACCAACCCACCTGCCAACGGGGCAAGTACTGCGAGGAGTGAAGGTGAAA 1683
Db 2705 GAGTGCAGGACCCGAGCCTCAGATGGAGGAAGATGCTGCTGGTGGAGAGCCAG 2764
QY 1684 TACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAG 1743

Db 2765 TACAGTCTATGCCACGAGGAATGCCCC-----CCTGACGGAAAGCTTCAGGGAG 2818
QY 1744 GAGCAGTGTGAGGCTTCAACGGCTACAAACACAGCACCACCGCTCACTCTCAAGCTG 1803
Db 2819 CAGCAGTGTGAGAGTAATAATGCCCTTACAATTACACTGAGGAGGGAATCT---CCTG 2875
QY 1804 GCATGGTGCCCAAGTACTCCGGGCTGTCTCCCGGGACAAGTGCAGCTCACTCAATGCGCA 1863
Db 2876 CAGTGGGTCCCAAGATATGCTGGGTGTCCCGGGACCGCTCAAGTGTCTGCGCA 2935
QY 1864 GCAATGACACTGGCTACTTCTATGTGTGACACCAAGGTGGTGGAGCGGACGTTGTC 1923
Db 2936 GCCCGGGGAGGAGCGAGTTCAAGTGTTCGAGGCAAGGTGATGATGACACCTTGTGT 2995
QY 1924 TCTCTGACTCCACTCCGCTGTGTCTCAAGCAAGTGCATCAAGGCTGGCTGTATGGG 1983
Db 2996 GGGCCAGAAACTGCGCATCTGTGTGCTGGCAGTGTGTCAAGCCGCTGTUACCAT 3055
QY 1984 AACCTGGGCTCCCAAGAAGAGATTTCGACAGTGTGGGTGTGTGGGGAGACAAATAGAGC 2043
Db 3056 GTGGTGGACTCGCTCGGAAGCTGGACAAATCGGGGTGTGTGGGGCAAGGCAACTCC 3115
QY 2044 TGAAGAAGTGAAGTCTTCAACAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2103
Db 3116 TGCAGGAAGTCTCCGGGTCCCTCACCCCACTATGCTGCTGCTGCTGCTGCTGCTGCT 3175
QY 2104 ATCCCGCAGGCGCTCAAGCATTCGACATCCGCGCAGCGGGTTACAAAGGGCTGATCGGG 2163
Db 3176 ATCCAGCTGTGTGCTCACTAATATGCTGAGCAGCGGAGCCACCGGCTGTGAGAAC 3235
QY 2164 GATGACAACTACTGCTCTGAAGAACACGCAAGCAAGTACTGCTCAACGGGCAATTC 2223
Db 3236 GATGGAACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3295
QY 2224 GTGGTGTGGGCTGAGCGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2283
Db 3296 GGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3355
QY 2284 ACGGGCACAGCGGTGGAGAGCCTTCCGAGGCTTCCCGCCCATCTGCTGCTGCTGCTG 2343
Db 3356 TCCATGCCACCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3415
QY 2344 GAGTCTCTTCCGT---GGGGAAGATGACACCCCGCGGCTGCTGCTGCTGCTGCTG 2400
Db 3416 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3475
QY 2401 CCCAAGA 2408
Db 3476 CCTAATGA 3483

RESULT 9

US-09-949-002-108
; Sequence 108, Application US/09949002
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 3711
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

Db 2893 GATGGAACTACCTGGCGCTGAGACGGCTGATGGCGAGTACCTGCTCAAGGCAACTG 2952
 QY 2224 GTGGTGTGGCGGCTGGAGCGGACCTGGTGGTGAAGGGGAGTCTGCTGCTGACAGCGG 2283
 Db 2953 GCCATCTCTGCTATAGACGAGGACATCTTGGTGAAGGGGACCATCTGAACTAGAGGG 3012
 QY 2284 ACGGGACAGCGGTGGAGAGCTGACGCTTCCCGGCCCATCTGGAGCGCTGACCGTG 2343
 Db 3013 TCCATCGCCACCTGGAGCGCTGACAGCTTCCGGCCCTTGGCAGAGCTCTGACAGTG 3072
 QY 2344 GAGTCTCTCCCGT---GGGGAAGATGACAGCGCCCGGGTCCGCTACTCTTCTATCTG 2400
 Db 3073 CAGCTCTGACAGTCCCTGGCGAGGCTTCCCGCCCAAAAGTCAATATACACTTCTTGT 3132
 QY 2401 CCCAAAGA 2408
 Db 3133 CCTATGA 3140

RESULT 10

US-09-724-676-18640
 ; Sequence 18640, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: CompuGen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 CompuGen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 18640
 ; LENGTH: 4016
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-18640

Query Match 23.3%; Score 664.4; Db 6; Length 4016;

Best Local Similarity 59.2%; Pred. No. 2e-133;
 Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

QY 157 CTCATTTTTCAGATCAGACGATTCAGAGGAGCTTTTACCTACACTGACGCGGATGCT 216
 Db 865 CTCGGCTCCACCTGTCGGCTTCGGCAAGGGCTTCGTGTGGAGCTGAGGCGCAGAC 924
 QY 217 CAGTTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGGGTCCCGCTCCAGGGGCTACC 276
 Db 925 AGCTTCTTGGCGCGGAGTTCAAGATCGAGCGCTCGGGGCTCCGGCGGCG---ACC 981
 QY 277 GGGGCTCTTCAGACCTCGGAGCTGCTTCTATTCTGGGGAGTGAACGCGAGCGGAC 336
 Db 982 GGGGCGAGCGGGGCTCGCGGCTGTTTTTTTTCGGGACCGTGAATGGGGGCGCGAG 1041
 QY 337 TCGTTCTGCTGTGAGCTGTGCGGGGCTCCCGGAGCGCTTGGCTACCGAGCGCC 396
 Db 1042 TCGCTGGCGGGCTCAGCTGTGCGGGGCTGAGCGGCTCTCTCTGCTGACGCGGAG 1101
 QY 397 GAGTATGTCATTAGCGCGCTGCCAATGCTAGCGCGCGCGCGAGCGCAACAGCGAG 456
 Db 1102 GAGTTACCATTCACCGCAGGCGCGGGGCTCCTGCTCAGCGCGCACCGCTGAG 1161
 QY 457 GCGCACACCTTCTCCAGCGCGGGGTGTTCCGGGCGGGCTTCCGAGACCCCACTCT 516
 Db 1162 CGCTGGGTCCCGGAGCGCGCGCTCCCGGAGGACCGAGTGGGAGTGGAGAGC 1221
 QY 517 CGCTCGGGGTG-GCTCGGCTGGAACCGCGCATCTTACGGCCCTGACCCCTTACAA 575
 Db 1222 GGAGAGGTTCAGAGCAGAGGAGAGGAGACACACGAGAGCAGAGGAGAGAGCCAA 1281
 QY 576 GCGCGGGGCGGGCTTCGGGGAGAGTCTGAGCGCGCAGG-----TCTGGGCGGCC 630
 Db 1282 GAAGAGGAGGAGAGGCGCTAGCGAGCGCGCGCGCTGGGGCGCACAGTAGGAGC 1341
 QY 631 AAGCGTTTCGTCTATCCCGCGTACGTGGAGACGCTGCTGGTCCGCGACGAGTCAATG 690

Db 1342 AAGCGTTTCGTCTGAGCGCGCTTCTGGTGGAGACGCTGCTGGTGGCGATGCTGCCATG 1401
 QY 691 GTCAGTTTCACGCGCGGACCTTGGGAACATATCTCTGCTGACGCTGCTGGCAACGCGCGG 750
 Db 1402 GCTGCTTCTACGCGCGGCGACCTTGCAGAACCATCTCTGACGTTAATGTTCTGTGCGCC 1461
 QY 751 CGACTCTACCGCCATCCAGCATCTCAAGCCCATCAACATCTTGTGTGTCGAAGTGGCTG 810
 Db 1462 CGATTCTACAGACGACCCAGCATCAAGATTCATCACTGATGCTGTGTAAGTGTGCTG 1521
 QY 811 CTCTTAGAGATCTGACTCGGGCCCAAGTCCAGCGGAATCCGCGCTCTGACGCTGGCG 870
 Db 1522 ATCGTAGAAGTGAATAATGGGCGCCAGAGTGTGTCGACAATGGGGGGTCTACACTGGT 1581
 QY 871 AACTTCTGCTGCTGGCAGAGAAGTGAACAAAGTGAAGTGAACAGCACCCTGACTGG 930
 Db 1582 AACTTCTGCAACTGGCAGGCGGTTTCAACAGCCAGGACGCGCACCCAGAGTACTAC 1641
 QY 931 GACTGTCATCTCTTCCAGCAGCAGGACCTGTGG---AGCCACACCTCTCACACC 987
 Db 1642 GACAGCGCATCTCTGCTACCCAGACAGAACTTCTGTGGGCGAGGAGGGCTGTGTGACACC 1701
 QY 988 CTGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAAGTCTCTCTGCTCATTTAG 1047
 Db 1702 CTGGCTGTGGCAGACATCGGGACCATTTGTGACCCCAACAAAGTCTCTCGCTGATCGAG 1761
 QY 1048 GACGATGGCTTCCATCAGCTTCCAGCTTCCAGCTGCGCCAGAGCTGGCCAGCTGTTCACATG 1107
 Db 1762 GATGAGGGGCTCCAGGCGGCCACCTGGCCCATGAATAGGCGACCTCTCTGAGCATG 1821
 QY 1108 CCCCATGACAAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCCCAACACATG 1167
 Db 1822 CCCCAGAGGACTTCAAGCCCTGACACGCGCTCTTCGGGCGCATGGGCAAGCACCAGTG 1881
 QY 1168 ATGTCGCGACCTTCATCCAGATCGACGCTGCGCAACCCCTGGTGCAGCTGCACTGGTGGC 1227
 Db 1882 ATGGCAGCGCTGTTCGTCCACTGAACCCAGACGCTGCGCTGGTCCCTCTGACGCGCATG 1941
 QY 1228 ATCATCACCGACTTCTTGACAGCGGCGAGGTGACTGCTCTCTGACCAACCCAGCAAG 1287
 Db 1942 TATCTCACAGAGCTTCTGGACGGCGGCGAGAGACTGTCTCTGATGCCCTGCTGGCG 2001
 QY 1288 CCACTCTCCCTCCCGAGATCTGCGGGC-----GCCAGCTACACCTTGACGAGCAG 1341
 Db 2002 GCGCTGCGCCCTCCCGAGGCTCCCGGCGCGCATGGCCCTGTACCACTGGGAGCAGCAG 2061
 QY 1342 TCGAGCTGGCTTTTGGGCTGGGTCCCAAGCCCTGTCTTACA-----TGCAGTAC 1392
 Db 2062 TGCAGGAGATCTTGGCGCGGATTTCCGCCACTGCCCCAACACCTCTGCTCAGGAGCTC 2121
 QY 1393 TGCACCAAGCTGTGTG---CACCGGAGGCGCAAGGACAGATGTTGTGCCAGACCCG 1449
 Db 2122 TCGGCGCAGCTTGTGTGCCACACTGATGGGCTGAGCCCTGTGCCACAGAAATGGC 2181
 QY 1450 CACTTCCCTGCGCGATGGCACCAGCTGTGGCGAGGCGCAAGCTCTGCTCAAGGGGCG 1509
 Db 2182 AGCTTCCCTGCGCTGACGCGAGCGCTGCGGGCTGGGCACTCTGCTCAGAAGCAGC 2241
 QY 1510 TG-----CGTGGAGACACACCTCAAGACGACAGGTTGGATGTTCTCTGGCCCAA 1563
 Db 2242 TGTCTACCTGAGGAGGAGTGGAGGCGCCCAAGCCCTGTGTAGTAGGAGCTGGGCGCG 2301
 QY 1564 TGGGATCCCTATGCGCCCTGCTCGCGACATGTGTGGTGGGCGCTGCACTGGCGCAGAG 1623
 Db 2302 TGGGAGCTTGGGAGATGTTCTCGGACCTGTGAGGAGGAGTACACTTTTTCACACCT 2361
 QY 1624 CAGTGCACCAACCCCTGCGCAACGGGGGCAAGTACTCTGCGAGGAGTGGAGGTTGAAA 1683
 Db 2362 GAGTGCAGAGGAGCGGAGCTCAGAAATGGAGGAAGTACTGCTGGGTGGGAGGAGCAG 2421
 QY 1684 TACCGATCTCTCAATCTGGAGCCCTGCGCCAGCTCAGCTCCGGAAGAGCTTCCGGGAG 1743

Db 2422 TACCAGTCATGCCACAGGGAATGCCCC-----CCTGACGGGAAAGCTTCAGGGAG 2475
Qy 1744 GAGCAGTGTGAGGCTTTCACGGCTACACACACACACACACACCGGCTCACTCTCGCGCGTG 1803
Db 2476 CAGCAGTGTGAGGAAGTATAAGCTTACAATTACATCGATGAGGAGGGAATCT---CCTG 2532
Qy 1804 GCATGGGTGCCAAGTACTCCGCGGTCTCCCGGGGACAAAGTGAAGCTCATCTGCGGA 1863
Db 2533 CAGTGGGTCCCAAGTATGCTGGGTGTCCCGCGGACCGCTGCAAGTGTCTGCGGA 2592
Qy 1864 GCCAATGGCACTGGCTACTCTATGTGTGCGCACCCACAGGTGGTGGAGGSCACGCTGTGC 1923
Db 2593 GCCCGGGGAGGAGCGAGTTCAAAGTGTTCGAGGCCCAAGGTGATGTAGTGGCAACCTGTGT 2652
Qy 1924 TCTCCTGACTCCACCTCCGCTCTGTGTGCCAAGCAAGTGCATCAAGGCTGGCTGTATGGG 1983
Db 2653 GGGCCAGAAACACTGGCCACTGTGTGCTGCGCCAGTGTGTCAAGGCGGCTGTGACCAT 2712
Qy 1984 AACCTGGGCTCCAAGAAGAGATTGCAAGTGTGGGTGTGTGGGGAGACATAAGACG 2043
Db 2713 GTGTGACTCGGCTCGGAAGCTGGACAATGCGGGGTGTGTGGGGCAAGGCAACTCC 2772
Qy 2044 TGCAAGAAGTGTGACTGTCTTCCAAAGCCCATGCTATGCTACAATTTGTTGGTGGCC 2103
Db 2773 TGCAGGAAGTCTCCGGGTCCCTACCCCAACCAATTTGCTACAAATGTGACATTTGCACC 2832
Qy 2104 ATCCCCGAGGCGCTCAAGCATGCAATCCGCGAGCGCGTTTACAAAGGCTGATCGGG 2163
Db 2833 ATCCCCGCTGTGGCACTAATTGAGTGAAGCAGCGAGGCCACCGGGTGTGCAGAAC 2892
Qy 2164 GATGACAATCTGCTGTGGAAGACGCCAAGCAAGTACCTGCTCAACGGGCATTTTC 2223
Db 2893 GATGGGAATCTGCTGCGCTGGAAGCGGCTGATGGCGAGTACCTGCTCAACGGCACTG 2952
Qy 2224 GTGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGCGCAGTCTGCTGCGCTACAGCGGC 2283
Db 2953 GCCATCTCTGCATAGACGAGGACATCTGTGTGAAGGGAACCATCTGAACTACAGCGGC 3012
Qy 2284 ACGGGACAGCGTGGAGAGCTGCAAGCTTCCCGGCCCATCTTGGAGCGCTGACCGGTG 2343
Db 3013 TCCATGCCCACTTGGAGCGCTGCAAGCTTCCGGCCCTTGGCAGAGCTCTGACAGCT 3072
Qy 2344 GAGTCTCTCCGT---GGGGAAGATGACACCGCGCCCGGCTCCCTACTCTCTATCTG 2400
Db 3073 CAGCTCTGACAGTCCCTGCGAGGCTTCCGCCCAAAAGTCAAAATACACCTTCTTGT 3132
Qy 2401 CCCAACA 2408
Db 3133 CCTATGA 3140

RESULT 11
US-09-724-676A-18640
; Sequence 18640, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18640
; LENGTH: 4016
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-18640

Query Match 23.3%; Score 664.4; DB 6; Length 4016;
Best Local Similarity 59.2%; Pred. No. 2e-133;
Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;
Qy 157 CTCATTTTTCAGATCAGACATTCAGAGGAGCTTTTACCTACACCTGACGCGCGGATGCT 216

Db 865 CTGCGGCTCACCTGTCCGCCCTTCGGCAAGAGGCTTCGTGTTCGCCCTTGGCGCCGCAAGAC 924
Qy 217 CAGTTCTTGGCTCCCGCCCTTCTCCACTGAGCATCTGGGCGGTCCCGCTCCAGGAGCTCAC 276
Db 925 AGCTTCTTGGCGCGCGCTTCAAGATCGAGGCGCTTCGGGGGCTCCGCGCGGGGCG---AGC 981
Qy 277 GGGGGCTTTCAGACCTGCGAGCTGCTTCTATTTCTGSGGACGTGAACGCGGAGCGGAC 336
Db 982 GGGGGGAGCGGGGCTGCGCGCTGTCTTTTTCGGCACCGTGAATGGGAGCCGAG 1041
Qy 337 TCGTTCGCTGTGTAGCCCTGTGCGGGGGCTTCGCGGAGCCTTTTGGCTTACCAGAGCGCC 396
Db 1042 TCGTTCGCGCGCTGAGCCTGTGCGCGGCTGAGCGGCTCTCTTCTGCTGAGCGCGAG 1101
Qy 397 GAGTATGTCTATTAGCCCGCTGCGCAATGTACGCGCGCGCGGCGAGCGACGCAACAGCCAG 456
Db 1102 GAGTTCAACATCCAGCGCGAGGCGCGGGGCTCCCTGGCTCAGCGCACCGCTGCAG 1161
Qy 457 GCGGCACACCTTCTCAGCGCGCGGGGTTCGCGGGCGGCTTCGCGGAGACCCACCTCT 516
Db 1162 CGCTGGGGTCCCGCGAGCGCCCTTCCCGGAGACCCCGAGTGGGAGGTGGAGACG 1221
Qy 517 CGCTGGGGGTG-GCCTCGGGCTGGAAACCGCGCATCTACGGCCCTTGACCCCTTACAA 575
Db 1222 GGAGAGGCTCAGAGGCGAGAGAGACACACAGGAGGACAGCGAGGAGGAGACCAA 1281
Qy 576 GCCGCGCGGGCGGCTTCGGGGAGAGTCTGAGCGCGCAGG-----TCTGGCGCGCC 630
Db 1282 GAGAGAGGAGAGAGGCGCTAGCGAGCGCGCACCGCCCTTGGGGGCCACGAGTAGAGCC 1341
Qy 631 AAGCGTTTCGTGTATATCCCGGCTACGTGGAGACGCTGTGTGGTGGCGGACGATCAATG 690
Db 1342 AAGCGGTTTGTCTGAGCGCGCTTCGTGGAGAGCTGTGTGGTGGCGGCTGCTGCTCATG 1401
Qy 691 GTCAAGTTCACGGCGGACCTGCAACATATCTGTGACGCTGTGCGCAACGCGCGGG 750
Db 1402 GCTGCGCTTCTACGGGCGACCTGCGAAGACCATCTCTGACGTTAATGTCTGTGCGAGCC 1461
Qy 751 CGACTCTACGGCATCCAGCATCTCAACCCCATCAACATCTGTGTGGTTCAGTGTCTG 810
Db 1462 CGATTCTCAAGCAACCCAGCATCAAGAAATTCATCAACCTGATGTTGGTGAAGTCTG 1521
Qy 811 CTCTTTAGAGATCGTGACTCCGGGCGCAAGGTCAACCGCAATGCGGCTTGAAGCTCCGC 870
Db 1522 ATGTTAGAGATGARAATGGGCGCCAGAGGTGTCCGACATGGGGGCTTACACTGCGT 1581
Qy 871 AACTTCTGCTGCGCAGAGAGAGCTGAACAAAGTGAAGTGAACAGACCCCGAGTACTGG 930
Db 1582 AACTTCTCAACTGGCAGCGCGCTTTCACACGCGCCAGCGGACCCACCCAGAGCACTAC 1641
Qy 931 GACACTGCCATCTCTTACACGAGGAGGACCTGTGTGG---AGCCACCACTGTGACAC 987
Db 1642 GACACGCCATCTCTGCTCACACAGACAGAACTTCTGTGGCAGGAGGGGTGTGTGACACC 1701
Qy 988 CTGGGATGGCTGATGTGGGTACCATGTGTACCCCAAGAGAGAGCTGCTGTCTATTGAG 1047
Db 1702 CTGGGTGTGGCAGACATCGGGACCATTTGTACCCCAACAAAGTCTGCTCGGTATCGAG 1761
Qy 1048 GAGCATGGGCTTCCATCAGCCTTCAACCATGCGGCGGAGCTGGGCGGACGTTTCAACATG 1107
Db 1762 GATGAGGGGCTCCAGCGCGCCACACCCCTGGCCCATGAACCTAGGGCACGTCCTTAGCATG 1821
Qy 1108 CCCCATGACAAATGTGAAGTCTGTGAGGAGGTGTGAGGAGCTCCGAGCGCAACACCATG 1167
Db 1822 CCCACAGCAGCTCAAGCCCTGACACAGGCTCTTCGCGGCCCATGGCAAGCAGCACGAGT 1881
Qy 1168 ATGTCCCGGACCCATCATCCAGATCCAGCGTGCACACCCCTGGTCAGCTTGCATGTGGCC 1227
Db 1882 ATGCAACCGCTGTGCTGCTACCTGAACCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1941
Qy 1228 ATCATCACCGACTTCTGTGACAGCGGCGAGCGTGAAGTCTGCTTCTGAGACCAACCAAG 1287

Db 1942 TATCTCAGAGACTTCTGGAGCGGGGACGAGAGACTGTCTCC TGGATGCCCCCTGGTGCG 2001
 Qy 1298 CCATCTCCCTCCGCGAGAGATCTGCGGGC-----GCCAGCTACACCCCTGAGCGACAG 1341
 Db 2002 GCGCTGCCCTCCGCGAGAGCTCCGCGGCGCATGCGCCCTGTACACAGCTGGACGACAG 2061
 Qy 1342 TGGAGCTGGCTTTTGGCGTGGGCTCCCAAGCCCTGTCTCTTACA-----TGCAGTAC 1392
 Db 2062 TGCAGGAGATCTTTGGGCGGATTTCCGCCACTGCCCAACACACTCTGTCTCAGGAGTC 2121
 Qy 1393 TGCACCAAGCTGTGGT---CACCGGAGGCCAAGGAAGATGCTGTGTCAGACCCGC 1449
 Db 2122 TGCGCCAGCTTTGGTGCCACACTGATGGGCGTGAGCCCTGTGCCACACCAAGATGGC 2181
 Qy 1450 CACTTCCCTCGGCGCATGGCCACAGCTGTGGCGAGGCAAGCTCTGCCTCAAGGGGCG 1509
 Db 2182 AGCTGCGCTGGCTGACGCGACGCGCTGCGGGCCTGGGCACTCTGTCTGAGAAGGCGC 2241
 Qy 1510 TG-----CCTGAGAGACACAACTCAACAGACAGAGGTGGATGTTCTTGGCCAAA 1563
 Db 2242 TGTCTACTGAGGAGAGTGGAGGCCCAAGCCCGTGTAGATGGAGGCTGGGCACTG 2301
 Qy 1564 TGGATCCCTATGGCCCTCTGCTGCGCACATGTGTGGGGGCGTGCAGCTGAGCGAGAG 1623
 Db 2302 TGGGCGCTGGGAGAAATGTTCTCGGACCTGTGGAGGAGGATACAGTTTTCACACCT 2361
 Qy 1624 CAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACIGCGAGGGAGTCAGGGTAAA 1683
 Db 2362 GAGTGCAGAGCCCGGAGCTCAGATGGAGGAAGATACGCTGGGTGCGAGAGCCAAAG 2421
 Qy 1684 TACGATCTCTGCAATCTGAGCCCTGCGCCAGCTCAGCGTCCGGAAGAGCTTCGCGSAG 1743
 Db 2422 TACCAGTATGCCACAGGAGGAATGCCCC-----CCTGACAGGAAAGCTTCAGGGAG 2475
 Qy 1744 GAGCAGTGTGAGGCTTCAAGGCTACACACAGCACCAACCGCTCACTCTCCGCGTG 1803
 Db 2476 CAGCAGTGTGAGAGTATATGCTTCAATATGCTGATGAGGAGGGAATCT---CCTG 2532
 Qy 1804 GCATGGGTGCCAAGTACTCCGCGTGTCTCCCGGGGACAACTGCAAGCTCAATCTGCCGA 1863
 Db 2533 CAGTGGGTGCCAAGTATGCTGGGTGTCCCGCGGGAGCGCTCAAGTGTTCCTGCCGA 2592
 Qy 1864 GCCAATGGCACTGCTACTTCTATGTGCTGGCACCAAGTGTGGGGGAGACAAAGAGC 1923
 Db 2593 GCCCGGGGAGGAGCGATCAAGTGTTCGAGGCCAAGCTGCAATGATGACACCTGTGT 2652
 Qy 1924 TCTCCTGACTCCAGCTCTGTGTCTCAAGGCAAGTGCATCAAGGCTGGCTGTGATGGG 1983
 Db 2653 GGGCCAGAAACACTGGCCATCTGTCTCGTGGCCAGTGTGTCAAGGCGGCTGTGACCAT 2712
 Qy 1984 AACCTGGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGGAGACAAAGAGC 2043
 Db 2713 GTGGTGGACTCGCTCGGAAGCTGGACAAATGCGGGGTGTGTGGGGCAAGGCACTCC 2772
 Qy 2044 TGCAGAGAGTGTGACTGTGACTTTCACCAAGCCCATGCTATGGCTACAAATTTCTGGTGGCC 2103
 Db 2773 TGCAGGAGTGTCTCGGGTCCCTTCAACCCACCAATTTATGCTACAAATGATGTCAAC 2832
 Qy 2104 ATCCCGCAGCGGCTCAAGCATCGACATCCGCGAGCGGCTTACAAAGGCGCTATGGG 2163
 Db 2833 ATCCAGCTGTGCCACTATATATGAGTGAAGCAGCGGAGGCCACCGGGTGTGCAGAAC 2892
 Qy 2164 GATGCAACTACTGTGGCTGTGAAGACACCAAGGCAAGTACCTGTCTCAATGGGCAATTC 2223
 Db 2893 GATGGGAACACTCTGGCGCTGAAGAGCGGCTGATGGGCGAGTACCTGTCTCAAGGCAACTG 2952
 Qy 2224 GTGGTGTGGCGGTGAGGAGGAGCTGTGGTGAAGGGCAAGTCTGCTCGCTACAGCGGC 2283
 Db 2953 GCCATCTCTGCCCATAGACAGGACATCTTGGTGAAGGGGACCACTCTGAAGTACAGCGGC 3012
 Qy 2284 AGGGCACAGCGGTGAGAGACCTTGAGGCTTCCGGCCCATCTCGAGCCCTGACCGTG 2343
 Db 3013 TCCATCGCCACCTTGAGCGCCTGACAGAGCTTCCGGCCCTTGCCAGAGGCTCTGACAGTG 3072

Qy 2344 GAGTCTCTCTCCGT---GGGAGAGATGACACGCGCGGGTCCGCTACTCTTCTATCTG 2400
 Db 3073 CAGCTCTCAGACTCCTGCGAGGCTTCCGCCCAAAAGTCAAAATACACTTCTTGTGT 3132
 Qy 2401 CCCAAAGA 2408
 Db 3133 CCTAATGA 3140

RESULT 12
 US-09-949-002-248
 ; Sequence 248, Application US/09949002
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
 ; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
 ; FILE REFERENCE: CL000790
 ; CURRENT APPLICATION NUMBER: US/09/949,002
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: 60/231,401
 ; NUMBER OF SEQ ID NOS: 10823
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 248
 ; LENGTH: 2664
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-002-248

Query Match 23.2%; Score 662.8; DB 6; Length 2664;
 Best Local Similarity 59.1%; Pred. NO. 4.1e-133;
 Matches 1353; Conservative 0; Mismatches 887; Indels 48; Gaps 11;

Qy 157 CTCATTTTTCAGATCAGACATTTTCAGGAGGACTTTTACCTACACCTGAGCGCGGATGCT 216
 Db 151 CTCGCGCTCCACCTCTCCCGCTTCGCGAAGGCTTCGTGCTGCGCTTGGCGCCGCGAGCAG 210
 Qy 217 CAGTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGCGCTCCCGCTTCCAGGAGCTCAAC 276
 Db 211 AGCTTCTTAGCGCCGAGTTCAGATCGAGCGCTTCGGGGGCTCCGGCGCGGGG---ACC 267
 Qy 277 GGGGCTCTTTCAGACCTGCGACGCTCTTCTTATCTGGGAGCTGAACGCGCGAGCGCGAC 336
 Db 268 GGGGCGAGCGGGGTGCGCGCTGCTTCTTCTCCGACCGCTGAATGGGAGCGCGAG 327
 Qy 337 TCGTTCGCTGCTGTGAGCCTGTGCGGGGGGCTTCCCGGAGCGCTTTGGCTACCGAGCGGC 396
 Db 328 TCGTGGCGGCTGAGCCTGTGCGCGGGCTGAGCGCTCTCTTCTGCTGAGCGCGAG 387
 Qy 397 GAGTATCTCATTAGCCCGCTGCCCAATGCTAGCGCGCGGGGCGAGCGCAACAGCGAG 456
 Db 388 GAGTTCACCATCAGCGCGAGCGGGGCTCCCTTGGCTCAGCGCGCACCGCTGCGAG 447
 Qy 457 GGGCCACACCTTCTCAGCGCGCGGGGTGTTCGGGCGGGGCTTCCGGAGAGACCACTCT 516
 Db 448 CGCTGGGGTCCCGCGAGCGCGCCCTTCCCGGAGGACCCGAGTGGGAGGTGAGAGCG 507
 Qy 517 CGCTGCGGGGTG---GCCTCGGGCTTGAACCCCGCCATCTACGGGCGCTTGGACCTTACAA 575
 Db 508 GGAGAGGTCAGAGGAGGAGAGAGAGAGACACACAGGAGGAGGAGGAGGAGAGCAAA 567
 Qy 576 GCCCGCGGGGGGCTTGGGGAGAGTGTAGCGCGGCGCAGG-----TCGCGCGCGC 630
 Db 568 GAAGAGGAGGAGAGAGCGCTAGCGAGCGCGCACCGCCCTGGGGGGCCACAGTATGAGCC 627
 Qy 631 AAGCGTTTCTGTCTATCCCGGCTACGTGAGAGCGCTGGTGGTGGGAGAGAGTCAATG 690
 Db 628 AAGCGGTTTGTGCTGAGGCGGCTTCTGAGAGCGCTGCTGGTGGCGGAGTGTCTCCATG 687
 Qy 691 GTCAAGTTTCACCGCGCGGAGCTTGGAGCAATATCTGCTGACGCTGCTGGCAACCGCGCG 750

```
Db 688 GCTGCTTTACTAGGGGGGCGACCTGCAGAACACATCCTGACGTTAATGCTGTGGCCAGCC 747
QY 751 CCACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCGTGTGCTGGTCAAGGTGCTG 810
Db 748 CGAATCTACAGCACCACGACATCAAGATTCATCAACATGATGTTGTTGAAAGTCTG 807
QY 811 CTTCTTAGAGATCGTACTCCGGGCCCCAGGTCAAGGCAATGCGGCCTCAACGCTCGCC 870
Db 808 ATCGTAGAAGATGAAATAGGGGCCAGAGGTTCGACAAATGAGGGGCTTACACTCGT 867
QY 871 AACTTCTGCTGCTGGCAGAGAGCTGACAAAGTGAAGTGAACAGACCCCGAGTACTGG 930
Db 868 AACTTCTGCAACTGGCAGCGGCTTCAACACGAGCCAGGACC3CCACCCAGACACTAC 927
QY 931 GACATGSCATCTCTTCCACGCGCAGGAGCTGTGG---AGCCACCACTGTGACACC 987
Db 928 GACAGGGCATCTCTGCTCAACAGACAGACTCTGTGGCAGAGGGCTGTGTGACACC 987
QY 988 CTGGCATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGTGTCTGTCTATTGAG 1047
Db 988 CTGGGTGTGGCAGACATCGGGGACCATTTGTGACCCCAACAAAGTGTCTGCTGATGAG 1047
QY 1048 GACGATGGGCTTCCATCAGCTTCCACACTGCCCCAGAGCTGGCGACGTGTCAACATG 1107
Db 1048 GATGAGGSCCTCAGGCGGCCACACCTGGCCCATGACTGACGACACTCTCTCAGCATG 1107
QY 1108 CCCCATGACATGTGAAGTCTGTGAGGAGTGTGTGGAGCTCGGAGCCCAACCATG 1167
Db 1108 CCCACAGAGACTTCCAGCCCTGCACACGCTCTTCGGGCCCATGGGCAACACACCATG 1167
QY 1168 ATGTCCCGACCCCTCATCCAGATGACCGTGTCCAAACCCCTGTCTGAGCTGAGTGTGCC 1227
Db 1168 ATGGCAGCGCTGTCTGCTCCACTGAACACAGACGCTGCGCTGTCTGCTGAGTGTGCC 1227
QY 1228 ATCATCAGCAGTCTCTGAGCAGGCGGAGTGTGCTGCTCTGACCAACACCCAGCAAG 1287
Db 1228 TATCTCAGACAGTCTCTGAGCAGGCGGAGTGTGCTGCTGAGTGTGCTGAGTGTGCC 1287
QY 1288 CCCATCTCTGCGGAGGATGTGCGGGC-----GCCAGTACACCTGTGAGCAGGAG 1341
Db 1288 GCGCTGCGCTTCCCGACAGGCTTCCGGGGCGCATGGCGCTGTACCACTGAGCAGCAG 1347
QY 1342 TCGAGCTGCTGCTTTGGCGTGGCTCCAAAGCCCTGTCTTACA-----TGCACTAC 1392
Db 1348 TGCAGCAGATCTTTGGCGGATTTCCGCGACTGCGCCCAACACCTCTGCTCAGACGTC 1407
QY 1393 TGCACCAAGCTCTGTG---CACGGGAAGGCCAAGGACAGATGTTGTGCGCAGACCCG 1449
Db 1408 TGGCGCCAGCTTTGTGCTCAGCTGTGAGGCTGAGCCCTGTGCCACACGAAGATGGC 1467
QY 1450 CACTTCCCTGGCGGATGCGACAGCTGTGGCGAGGCAAGCTCTGCTCAAGGGGCC 1509
Db 1468 AGCTGCGCTGGCTGACGGCAGCGCTGCGGGCTTGGCAGCTCTGCTCAGAGGCGAGC 1527
QY 1510 TG-----CGTGGAGAGACAACTCAACAAGCAGGAGTGTGCTTCCGGCCAAA 1563
Db 1528 TGTCTACCTGAGGAGAGTGGAGAGGCCCAAGCCCGTGGCAGTGGAGGCTGGCCACG 1587
QY 1564 TGGGATCCATATGCCCCCTGCTCGCGCAATGTGTGGGGCGTGTGAGCTGGCCAGGAG 1623
Db 1588 TGGGACCCCTGGGAGAGATGTTCTCGAACCTGTGGAGGAGGATACAGTTTTCACACCGT 1647
QY 1624 CAGTGACCAACCCACCCCTGCGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGTGAAA 1683
Db 1648 GAGTGAGGAGCCCGAGCCCTCAGATGGAGGAAGATACTGCTGGTGGGAGGCCAAG 1707
QY 1684 TACCATCTGCAATCTTGAGCCCTGCGCAGCTGAGCTCCGGAAGAGCTTCCGGGAG 1743
Db 1708 TACCATGATGCGACAGGAGAAATGCCCC-----CTGACGGGAAGCTTCAAGGAG 1761
QY 1744 GAGCAGTGTGAGGCTTTAAGGGCTACAAACACAGACCAACCGCTCACTCTCGCCGTG 1803
Db 1762 CAGCAGTGTGAGAGTATAATGCTTACATTTACACTGACATGGAGGGGAACTCT---CCTG 1818
```

```
QY 1804 GCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAGTCAAGCTCATCTGCCGA 1863
Db 1819 CAGTGGTCCCAGATGATGTGGGTGTCTCCCGGGACCGCTGCAAGTTGTTCTGCCGA 1878
QY 1864 GCCAATGGCAGTGGCTACTTCTATGTGTGCACCCCAAGGTGTGGACGCCAGCTGTGTC 1923
Db 1879 GCCCGGGGAGGAGGAGTTCAAAGTTCGAGGCCAAGGTGATTGATGCCACCTGTGT 1938
QY 1924 TCTCCTGACTTCCACACTCCGCTGTGTCTCCAAAGCAAGTGCATCAAGGCTGCTGTATGGG 1983
Db 1939 GGGCAGAAACACTGGCCACTGTGTCTGCGCAGTGTGTCAAGGCCGCTGTGACCAT 1998
QY 1984 AACTGGGCTTCAAGAAGAGATTCCACAAGTGTGGGTGTGTGGGGAGACAAATAGAGC 2043
Db 1999 GTGGTGGACTCGCTCGGAGCTGCAAAATGCGGGGTGTGTGGGGCAAGCAACTCC 2058
QY 2044 TGCAGAGAGTGAAGTGTCTTCAACAGCCCATGATGATGCTACAAATTTCTGTGTGGCC 2103
Db 2059 TGCAGGAAGGTCTCGGGGTCCCTCACCCCCACCAATATTATGGCTACATGACAATGTCACC 2118
QY 2104 ATCCCCGACGCGCTCAAGCATCGACATCCGCGCAGCGGCTTACAAAGGCTGTATCGGG 2163
Db 2119 ATCCAGCTGGTCCCACTAATATTGACCTGAAGCAGCGGAGCCACCGGCTGTGAGAAC 2178
QY 2164 GATCACAACCTACTGCTCTGAAGAACAAGCAAGCAAGTACCTGTCAACGGGCAATTC 2223
Db 2179 GATGGGAACTACTGCGCTGANGACGCTGATGGCAGTACCTGTCTACAGGCAACTG 2238
QY 2224 GTGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGTGCGGTACAGCGGC 2283
Db 2239 GCCATCTCTGCCATAGCAGGACATCTTGTGAAGGGAGCCATCTTGAAGTACAGCGC 2298
QY 2284 ACGGGACACGCGTGGAGAGCCTGCGAGCTTCCGGGCCATCTTGGAGCGCTGACCGTG 2343
Db 2299 TCCATCGCACCTCGAGCGCTGCGAGCTTCCGGCCCTTGGCAGAGCTCTGACAGTG 2358
QY 2344 GAGTCTCTCTCGT---GGGGAAGATGACAGCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2359 CAGTCTCTGACAGTCTCTGCGAGGCTTCCCGCCCAAAAGTCAANTACACCTCTTTGT 2418
QY 2401 CCCAAGA 2408
Db 2419 CCTAATGA 2426
```

RESULT 13
US-10-170-235-34404
; Sequence 34404, Application US/10170235
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU

; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

; FILE REFERENCE: CL001380

; CURRENT APPLICATION NUMBER: US/10/170,235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 34404

; LENGTH: 3933

; TYPE: DNA

; ORGANISM: HUMAN

US-10-170-235-34404

Query Match 23.2%; Score 662.8; DB 8; Length 3933;

Best Local Similarity 59.1%; Pred. No. 4.5e-133;

Matches 1353; Conservative 0; Mismatches 887; Indels 48; Gaps 11;

QY 157 CTGATTTTTCAGATCAGACATTTTCAGGAGGACTTTTACCTACACCTGAGCGCGGATGCT 216

Db 1208 CTCGGCTCCACCTGTCCCGCTTCGGCAAGGCTTCGTGCTGCGCTTGGCGCCCGACAGAC 1267

QY 217 CAGTTTCTGGCTCCCGCTTCTCCACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACC 276

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:22:22 ; Search time 159 Seconds
(without alignments)
3852.179 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLILTAFAAGTAGGSE.....DQCNLHRKPELDPCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main :

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pap.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pap.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pap.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pap.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pap.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pap.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pap.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pap.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pap.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pap.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pap.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pap.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pap.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pap.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pap.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pap.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pap.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pap.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pap.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pap.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pap.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pap.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pap.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pap.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pap.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pap.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	5162	100.0	950	21	US-09-741-151-2	Sequence 2, Appli
2	5162	100.0	950	23	US-09-965-631-4	Sequence 4, Appli
3	5156	99.9	950	24	US-10-009-332-1	Sequence 1, Appli
4	5117	99.1	952	27	US-60-216-821-1	Sequence 4, Appli
5	5051	97.8	1132	27	US-60-212-656-497	Sequence 497, App
6	5051	97.8	1207	27	US-60-230-435-1155	Sequence 1155, Ap

7	5041	97.7	1252	27	US-60-242-679-1059	Sequence 1059, Ap
8	5038	97.6	1158	27	US-60-207-315-292	Sequence 292, App
9	4921	95.3	924	24	US-10-093-463-28	Sequence 28, Appl
10	4758	92.2	884	27	US-60-208-020-103	Sequence 103, App
11	4758	92.2	884	27	US-60-209-043-133	Sequence 133, App
12	4248.5	82.3	823	25	US-10-163-316-2	Sequence 2, Appli
13	4248.5	82.3	823	25	US-60-297-863-2	Sequence 2, Appli
14	3963.5	76.8	755	24	US-10-093-463-30	Sequence 30, Appl
15	3528	49.0	472	27	US-60-206-028-172	Sequence 172, App
16	2488	48.2	462	27	US-60-206-028-173	Sequence 173, App
17	2487.5	48.2	931	21	US-09-791-537-96880	Sequence 96880, A
18	2486.5	48.2	931	21	US-09-741-151-4	Sequence 4, Appli
19	2486.5	48.2	949	19	US-09-568-559-2	Sequence 2, Appli
20	2486.5	48.2	950	1	PCT-US00-14462A-2	Sequence 2, Appli
21	2486.5	48.2	950	1	PCT-US99-01313-2	Sequence 2, Appli
22	2486.5	48.2	950	16	US-09-235-810-2	Sequence 2, Appli
23	2486.5	48.2	950	17	US-09-373-658-2	Sequence 2, Appli
24	2486.5	48.2	950	17	US-09-373-658B-2	Sequence 2, Appli
25	2486.5	48.2	950	17	US-09-373-658C-2	Sequence 2, Appli
26	2486.5	48.2	950	21	US-09-791-537-108327	Sequence 108327, A
27	2486.5	48.2	950	23	US-09-989-687-2	Sequence 2, Appli
28	2486.5	48.2	950	27	US-60-171-503-2	Sequence 2, Appli
29	2486.5	48.2	967	12	US-08-845-456-2	Sequence 2, Appli
30	2486.5	48.2	967	17	US-09-388-279-2	Sequence 2, Appli
31	2486.5	48.2	967	17	US-09-388-280-2	Sequence 2, Appli
32	2486.5	48.2	967	21	US-09-791-537-41654	Sequence 41654, A
33	2486.5	48.2	967	24	US-10-057-084-2	Sequence 2, Appli
34	2486.5	48.2	967	25	US-10-105-929-2	Sequence 2, Appli
35	2486.5	48.2	967	25	US-10-115-286-2	Sequence 2, Appli
36	2486.5	48.2	968	1	PCT-US00-14462A-125	Sequence 125, App
37	2486.5	48.2	968	17	US-09-373-658B-125	Sequence 125, App
38	2486.5	48.2	968	17	US-09-373-658B-125	Sequence 125, App
39	2486.5	48.2	968	27	US-60-171-503-125	Sequence 125, App
40	2486.5	48.2	968	27	US-09-791-537-7136	Sequence 7136, App
41	2486.5	48.2	968	27	US-09-321-987B-4	Sequence 4, Appli
42	2486.5	48.1	951	16	US-09-264-585-17	Sequence 17, Appli
43	2485.5	48.1	968	25	US-10-163-316-7	Sequence 7, Appli
44	2485.5	48.1	968	27	US-60-297-863-7	Sequence 7, Appli
45	2482.5	48.1	967	21	US-09-791-537-49234	Sequence 49234, A

ALIGNMENTS

RESULT 1
US-09-741-151-2
; Sequence 2, Application US/09741151
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shiaoqing et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Human
US-09-741-151-2

Query Match 100.0%; Score 5162; DB 21; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLLGLILTAFAAGTAGGSEPVVPIRLDPDINGRYWRGPDSDGDLIFQIIAF	60
Db	1	MLLGLILTAFAAGTAGGSEPVVPIRLDPDINGRYWRGPDSDGDLIFQIIAF	60
QY	61	QEDFYHLTPDAQFLAPAFSTHGLVPLQSLTGSSDLRRCFYSGDVNAEPDSFAVSLC	120

```
Db 61 QEDFYHLTPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEPDFAAASLC 120
QY 121 GGLGAFGYGAEEYVISPFPNASAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180
Db 121 GGLGAFGYGAEEYVISPFPNASAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDYPKPRRAGFGESESRSSRRSGRAKRFVSIPRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDYPKPRRAGFGESESRSSRRSGRAKRFVSIPRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAMOKKLN 300
Db 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAMOKKLN 300
QY 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSQSVIEDDGLPSAFTTA 360
Db 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSQSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNMPHDNVKVCVEEFGKLRANHMSPITLQIDRANPWSACSAAITDFLDSHG 420
Db 361 HELGHVFNMPHDNVKVCVEEFGKLRANHMSPITLQIDRANPWSACSAAITDFLDSHG 420
QY 421 DCLLDQSPKPSISLPEDLPFGASYTISOQCELAFGVSKPCPYMOYCTKLWCTGKAKGOMVC 480
Db 421 DCLLDQSPKPSISLPEDLPFGASYTISOQCELAFGVSKPCPYMOYCTKLWCTGKAKGOMVC 480
QY 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKHRYVDGSMKAWDPYPCSRCTCGGVOLAR 540
Db 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKHRYVDGSMKAWDPYPCSRCTCGGVOLAR 540
QY 541 RQCTNPTPANGKYCEGVRYKYSCLNLEPCPSASGSKSPREEDQCEAFNGYNHSTNRLTIA 600
Db 541 RQCTNPTPANGKYCEGVRYKYSCLNLEPCPSASGSKSPREEDQCEAFNGYNHSTNRLTIA 600
QY 601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVYDGLTCLSPDSTSVYCVGKCIKAGCD 660
Db 601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVYDGLTCLSPDSTSVYCVGKCIKAGCD 660
QY 661 GNLGSKKRFDCGVCVGDNSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Db 661 GNLGSKKRFDCGVCVGDNSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
QY 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLT 780
Db 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLT 780
QY 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHHPKOPRGPVLSHNSVLSLSNOVEQDORPP 840
Db 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHHPKOPRGPVLSHNSVLSLSNOVEQDORPP 840
QY 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGQRTVPACDAAHRPVETQACGECPTWELS 900
Db 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGQRTVPACDAAHRPVETQACGECPTWELS 900
QY 901 AWSPCSKCGRGFORRSKLCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950
Db 901 AWSPCSKCGRGFORRSKLCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950
```

RESULT 2

```
US-09-965-631-4
; Sequence 4, Application US/09965631
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-965-631-4
```

Query Match 100.0%; Score 5162; DB 23; Length 950;

Best Local Similarity 100.0%; Pred. No. 0; Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGILTLTAFAGTAGGSEPEREVVPIRLDPDINGRYYIWRGPDGSGOGLIFQITAF 60

Db 1 MLLGILTLTAFAGTAGGSEPEREVVPIRLDPDINGRYYIWRGPDGSGOGLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEPDFAAASLC 120

Db 61 QEDFYHLTPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEPDFAAASLC 120

QY 121 GGLGAFGYGAEEYVISPFPNASAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180

Db 121 GGLGAFGYGAEEYVISPFPNASAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180

QY 181 NPAILRALDYPKPRRAGFGESESRSSRRSGRAKRFVSIPRYVETLVVADESVMKFGADLEH 240

Db 181 NPAILRALDYPKPRRAGFGESESRSSRRSGRAKRFVSIPRYVETLVVADESVMKFGADLEH 240

QY 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAMOKKLN 300

Db 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAMOKKLN 300

QY 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSQSVIEDDGLPSAFTTA 360

Db 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSQSVIEDDGLPSAFTTA 360

QY 361 HELGHVFNMPHDNVKVCVEEFGKLRANHMSPITLQIDRANPWSACSAAITDFLDSHG 420

Db 361 HELGHVFNMPHDNVKVCVEEFGKLRANHMSPITLQIDRANPWSACSAAITDFLDSHG 420

QY 421 DCLLDQSPKPSISLPEDLPFGASYTISOQCELAFGVSKPCPYMOYCTKLWCTGKAKGOMVC 480

Db 421 DCLLDQSPKPSISLPEDLPFGASYTISOQCELAFGVSKPCPYMOYCTKLWCTGKAKGOMVC 480

QY 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKHRYVDGSMKAWDPYPCSRCTCGGVOLAR 540

Db 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKHRYVDGSMKAWDPYPCSRCTCGGVOLAR 540

QY 541 RQCTNPTPANGKYCEGVRYKYSCLNLEPCPSASGSKSPREEDQCEAFNGYNHSTNRLTIA 600

Db 541 RQCTNPTPANGKYCEGVRYKYSCLNLEPCPSASGSKSPREEDQCEAFNGYNHSTNRLTIA 600

QY 601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVYDGLTCLSPDSTSVYCVGKCIKAGCD 660

Db 601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVYDGLTCLSPDSTSVYCVGKCIKAGCD 660

QY 661 GNLGSKKRFDCGVCVGDNSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720

Db 661 GNLGSKKRFDCGVCVGDNSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720

QY 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLT 780

Db 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLT 780

QY 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHHPKOPRGPVLSHNSVLSLSNOVEQDORPP 840

Db 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHHPKOPRGPVLSHNSVLSLSNOVEQDORPP 840

QY 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGQRTVPACDAAHRPVETQACGECPTWELS 900

Db 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGQRTVPACDAAHRPVETQACGECPTWELS 900

QY 901 AWSPCSKCGRGFORRSKLCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950

Db 901 AWSPCSKCGRGFORRSKLCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950

RESULT 3									
US-10-009-332-1									
; Sequence 1, Application US/10009332									
; GENERAL INFORMATION:									
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.									
; TITLE OF INVENTION: KAZUSA DNA RESEARCH INSTITUTE									
; FILE REFERENCE: Q67541									
; CURRENT APPLICATION NUMBER: US/10/009,332									
; PRIOR FILING DATE: 2001-12-10									
; PRIOR FILING DATE: 1999-11-11									
; PRIOR APPLICATION NUMBER: JPA 2000-144020									
; PRIOR FILING DATE: 2000-05-16									
; NUMBER OF SEQ ID NOS: 35									
; SOFTWARE: Patent version 3.1									
; SEQ ID NO 1									
; LENGTH: 950									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-009-332-1									
Query Match 99.9%; Score 5156; DB 24; Length 950;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	MLLGILTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF	60						
DB	1	MLLGILTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF	60						
QY	61	QEDFYHLTDPADFLAPAFSTEHLGVPLQGLTGGSSDLRCFYSGDVNAEPDFAAVSLC	120						
DB	61	QEDFYHLTDPADFLAPAFSTEHLGVPLQGLTGGSSDLRCFYSGDVNAEPDFAAVSLC	120						
QY	121	GGLRGAGYGAEVVISPLPNASAPAAQRNSOGAHLRLQRGVPGSGDPTSRGCVASGW	180						
DB	121	GGLRGAGYGAEVVISPLPNASAPAAQRNSOGAHLRLQRGVPGSGDPTSRGCVASGW	180						
QY	181	NPAILRALDPYKPRRAGGESRRSRGRKRFRVSIPIRYVETLTVVADSVKPHGADLEH	240						
DB	181	NPAILRALDPYKPRRAGGESRRSRGRKRFRVSIPIRYVETLTVVADSVKPHGADLEH	240						
QY	241	YLLTLLTAARLYRHPSTLNPINIVVVKVLLLRDRSGPKVTGNAATLNFCAWQKLN	300						
DB	241	YLLTLLTAARLYRHPSTLNPINIVVVKVLLLRDRSGPKVTGNAATLNFCAWQKLN	300						
QY	301	KVSDKHPEYWDTAILFTQDLGCGATTCDTLGMADVGTMCDPKRCSCVIEDDGLPSAFTTA	360						
DB	301	KVSDKHPEYWDTAILFTQDLGCGATTCDTLGMADVGTMCDPKRCSCVIEDDGLPSAFTTA	360						
QY	361	HELGHVFNMPHDNVKVCBEVFGKLRANHMSPITLIQIDRANPWSACSAAIITDLDGSHG	420						
DB	361	HELGHVFNMPHDNVKVCBEVFGKLRANHMSPITLIQIDRANPWSACSAAIITDLDGSHG	420						
QY	421	DCLLDQPSKPISLPDLPGASVYLSOQCELAGFVGSKEPCYMOYCTKLWCTGKAKGOMVC	480						
DB	421	DCLLDQPSKPISLPDLPGASVYLSOQCELAGFVGSKEPCYMOYCTKLWCTGKAKGOMVC	480						
QY	481	QTRHPFWADGTSCGSGKLCLGACVERHNLNKRHVDSWAKWDPYGPCSRITCGGGVQLAR	540						
DB	481	QTRHPFWADGTSCGSGKLCLGACVERHNLNKRHVDSWAKWDPYGPCSRITCGGGVQLAR	540						
QY	541	ROCTNPTPANGKYCEGVVRYKRSNLEPCPSSASGKSFIEQCEAFNGYNHSTNRLTLA	600						
DB	541	ROCTNPTPANGKYCEGVVRYKRSNLEPCPSSASGKSFIEQCEAFNGYNHSTNRLTLA	600						
QY	601	VAMVPKYSVSPRDKKILCRANGTGYFYVLAPKVVDGTCLGSPDSTVCVCGKCIKAGCD	660						
DB	601	VAMVPKYSVSPRDKKILCRANGTGYFYVLAPKVVDGTCLGSPDSTVCVCGKCIKAGCD	660						
QY	661	GNLGSKKRFDKGCGVCGDGNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720						

RESULT 4									
US-60-216-821-4									
; Sequence 4, Application US/60216821									
; GENERAL INFORMATION:									
; APPLICANT: Delegeane, Angelo M.									
; APPLICANT: Tang, Y. Tom									
; APPLICANT: Walsh, Roderick T.									
; APPLICANT: Lal, Preeti									
; APPLICANT: Khan, Farrah A.									
; APPLICANT: Nguyen, Dannie B.									
; APPLICANT: Tribouley, Catherine M.									
; APPLICANT: Yue, Henry									
; APPLICANT: Hafalia, April									
; APPLICANT: Patterson, Chandra									
; APPLICANT: Lu, Dying, Aina M.									
; APPLICANT: Azimzai, Yalda									
; APPLICANT: Gandhi, Ameena									
; APPLICANT: Baughn, Mariah R.									
; APPLICANT: Yao, Monique G.									
; APPLICANT: Lee, Ernestine A.									
; APPLICANT: Lu, Yan									
; APPLICANT: Ramkumar, Jaya									
; APPLICANT: Kallick, Deborah A.									
; APPLICANT: Au-Young, Janice									
; APPLICANT: Xu, Yuming									
; TITLE OF INVENTION: PROTEASES									
; FILE REFERENCE: PI-0146 P									
; CURRENT APPLICATION NUMBER: US/60/216,821									
; CURRENT FILING DATE: 2000-07-07									
; NUMBER OF SEQ ID NOS: 10									
; SOFTWARE: PERL Program									
; SEQ ID NO 4									
; LENGTH: 952									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
; NAME/KEY: misc.feature									
; OTHER INFORMATION: Incyte ID No: 7473089CD1									
US-60-216-821-4									
Query Match 99.1%; Score 5117; DB 27; Length 952;									
Best Local Similarity 99.5%; Pred. No. 0;									
Matches 947; Conservative 0; Mismatches 3; Indels 2; Gaps 2;									
QY	1	MLLGILTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF	60						
DB	1	MLLGILTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF	60						
QY	61	QEDFYHLTDPADFLAPAFSTEHLGVPLQGLTGGSSDLRCFYSGDVNAEPDFAAVSLC	120						
DB	61	QEDFYHLTDPADFLAPAFSTEHLGVPLQGLTGGSSDLRCFYSGDVNAEPDFAAVSLC	120						
QY	121	GGLRGAGYGAEVVISPLPNASAPAAQRNSOGAHLRLQRGVPGSGDPTSRGCVASGW	180						

Db 121 GGLRGAAGYRGAEYVISLPLNASAPAAQNSQGAHLQRRGVPGGSGDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGGEGESRRSRRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDPYKPRRAGGEGESRRSRRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLATAARLYRHPSTILNINIVVVKVLLRRDSDGPKVTGNAALTLENFCAWQKKLN 300
Db 241 YLLTLATAARLYRHPSTILNINIVVVKVLLRRDSDGPKVTGNAALTLENFCAWQKKLN 300
QY 301 KYSDKHPEYWDTAIFTTQDLGATTCDTLGMADVGMWCDPKRSCSVIEDDGLPSAFPTTA 360
Db 301 KYSDKHPEYWDTAIFTTQDLGATTCDTLGMADVGMWCDPKRSCSVIEDDGLPSAFPTTA 360
QY 361 HELGHVFNPHNDNVKVEEVFGKLRANHMSPPTLIQIDRANPWSACSAAIITDFLDSGGH 420
Db 361 HELGHVFNPHNDNVKVEEVFGKLRANHMSPPTLIQIDRANPWSACSAAIITDFLDSGGH 420
QY 421 DCLLDOPSKPISLBDLPASVYTLISOQCELAFAFGVSKPCPYMOYCTKLWCTGKAKGMVC 480
Db 421 DCLLDOPSKPISLBDLPASVYTLISOQCELAFAFGVSKPCPYMOYCTKLWCTGKAKGMVC 480
QY 481 QTRHFPWADGTCGSGKCLKLGACVERHNLNKRVDSWAKWDYGPCSRTCGGGVOLAR 540
Db 481 QTRHFPWADGTCGSGKCLKLGACVERHNLNKRVDSWAKWDYGPCSRTCGGGVOLAR 540
QY 541 ROCTNPTPANGKYGCEGVVRYRSCNLEPCPSSASGSKSFEEQCEAFNGYHNSTNRLTLA 600
Db 541 ROCTNPTPANGKYGCEGVVRYRSCNLEPCPSSASGSKSFEEQCEAFNGYHNSTNRLTLA 600
QY 601 VAWPKYSGVSPRDKCKLICRANGTGYYVLAPK-VVDGTLCSPDSTSVCVQKCKIKAGC 659
Db 601 VAWPKYSGVSPRDKCKLICRANGTGYYVLAPK-VVDGTLCSPDSTSVCVQKCKIKAGC 660
QY 660 DGNLGSKKRFPDKCGVCGDNKSKCKVTGLFTKPMHGFNFVVAIPAGASSIDIRQGYKGL 719
Db 661 DGNLGSKKRFPDKCGVCGDNKSKCKVTGLFTKPMHGFNFVVAIPAGASSIDIRQGYKGL 720
QY 720 IGDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILLEPL 779
Db 721 IGDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILLEPL 780
QY 780 TVEVLVSGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDR-GPSVLHNSVLNSNQVEQDDR 838
Db 781 TVEVLVSGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDR-GPSVLHNSVLNSNQVEQDDR 840
QY 839 PPARWAGSWGPCASCSGSLQKRAVDRCGAGORTVPACDAHRPVETQACGECPTWE 898
Db 841 PPARWAGSWGPCASCSGSLQKRAVDRCGAGORTVPACDAHRPVETQACGECPTWE 900
QY 899 LSAWSPCKSCGSGRQFRLSKCVGHGGRLLARDQCNLHRKPEQLDFCVLRPC 950
Db 901 LSAWSPCKSCGSGRQFRLSKCVGHGGRLLARDQCNLHRKPEQLDFCVLRPC 952

RESULT 5
US-60-212-656-497
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000673
; CURRENT APPLICATION NUMBER: US/60/212,656
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 497
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: HUMAN

US-60-212-656-497
Query Match 97.8%; Score 5051; DB 27; Length 1132;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 182; Gaps 2;
QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINRRRYWRCPEDSGDGLFIQTAF 60
Db 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINRRRYWRCPEDSGDGLFIQTAF 60
QY 61 QDEFLHLTPOAQFLAPAFSTEHLGVLPQLGTLGSSDLRRCFCYSGDVNAEPDFAAVSLC 120
Db 61 QDEFLHLTPOAQFLAPAFSTEHLGVLPQLGTLGSSDLRRCFCYSGDVNAEPDFAAVSLC 120
QY 121 GGLRGAAGYRGAEYVISLPLNASAPAAQNSQGAHLQRRGVPGGSGDPTSRGCVASGW 180
Db 121 GGLRGAAGYRGAEYVISLPLNASAPAAQNSQGAHLQRRGVPGGSGDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGGEGESRRSRRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDPYKPRRAGGEGESRRSRRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLATAARLYRHPSTILNINIVVVKVLLRRDSDGPKVTGNAALTLENFCAWQKKLN 300
Db 241 YLLTLATAARLYRHPSTILNINIVVVKVLLRRDSDGPKVTGNAALTLENFCAWQKKLN 300
QY 301 KYSDKHPEYWDTAIFTTQDLGATTCDTLGMADVGMWCDPKRSCSVIEDDGLPSAFPTTA 360
Db 301 KYSDKHPEYWDTAIFTTQDLGATTCDTLGMADVGMWCDPKRSCSVIEDDGLPSAFPTTA 360
QY 318 ----- 317
Db 318 ----- 317
QY 361 IRRRGAGTQEAAPPELOPVSKTDRRRREDGVGEGNGPPLLGRLLSLHLLRSPFKGCR 420
Db 361 IRRRGAGTQEAAPPELOPVSKTDRRRREDGVGEGNGPPLLGRLLSLHLLRSPFKGCR 420
QY 318 -----RGEL 321
Db 421 TEGDPRCSNPKLFYKKGQLENNSPGGLRFLAKMIPGSSGKVAATAKNSQLLIGVRCDL 480
QY 322 CGATTCTDGLMADVTMCDPKRSCSVIEDDGLPSAFSTAHELGHVFNPHNDNVKVEEVF 381
Db 481 CGATTCTDGLMADVTMCDPKRSCSVIEDDGLPSAFSTAHELGHVFNPHNDNVKVEEVF 540
QY 382 GKLRANHMSPPTLIQIDRANPWSACSAAITDFLDSGGDLDDPSKPSLSLEDLPFGAS 441
Db 541 GKLRANHMSPPTLIQIDRANPWSACSAAITDFLDSGGDLDDPSKPSLSLEDLPFGAS 600
QY 442 YTLSSQCELAFAFGVSKPCPYMOYCTKLWCTGKAKGMVCQTRHFPWADGTCGEGKLCIK 501
Db 601 YTLSSQCELAFAFGVSKPCPYMOYCTKLWCTGKAKGMVCQTRHFPWADGTCGEGKLCIK 660
QY 502 GACVERHNLNKHR-----VDGSAKWWDYPGCSRTCGGVQL 538
Db 661 GACVERHNLNKHRALTDIISPKQLLLRPNGLHTTKYDGSNAKWDPYGPCSRTCGGVQL 720
QY 539 ARROCTNPTPANGKYGCEGVVRYRSCNLEPCPSSASGSKSFEEQCEAFNGYHNSTNFLT 598
Db 721 ARROCTNPTPANGKYGCEGVVRYRSCNLEPCPSSASGSKSFEEQCEAFNGYHNSTNFLT 780
QY 599 LAVAWPKYSGVSPRDKCKLICRANGTGYYVLAPKVVVDGTLCSPDSTSVCVQKCKIKAG 658
Db 781 LAVAWPKYSGVSPRDKCKLICRANGTGYYVLAPKVVVDGTLCSPDSTSVCVQKCKIKAG 840
QY 659 CDGNLGSKKRFPDKCGVCGDNKSKCKVTGLFTKPMHGFNFVVAIPAGASSIDIRQGYKGL 718
Db 841 CDGNLGSKKRFPDKCGVCGDNKSKCKVTGLFTKPMHGFNFVVAIPAGASSIDIRQGYKGL 900
QY 719 LIGDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILLEP 778
Db 901 LIGDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILLEP 960
QY 779 LTVEVLVSGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDRGPSVLHNSVLNSNQVEQDDR 838
Db 961 LTVEVLVSGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDRGPSVLHNSVLNSNQVEQDDR 1020

Qy 839 PPARWAGSWGPGSCGSLGKRAVDRCGSAGQRTVPACDAHRPVTACGECPTWE 998
Db 1021 PPARWAGSWGPGSCGSLGKRAVDRCGSAGQRTVPACDAHRPVTACGECPTWE 1080
Qy 899 LSWSPCSKSCGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 950
Db 1081 LSWSPCSKSCGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 1132

RESULT 6
US-60-230-435-1155
; Sequence 1155, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US/60-230-435
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60-230-435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1155

Query Match 97.8%; Score 5051; DB 27; Length 1207;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 182; Gaps 2;

Qy 1 MLLGILTLTAFAGTAGGSEPERVVPRLDPDINGRRYWRGPDGSGGLIFQTAF 60
Db 76 MLLGILTLTAFAGTAGGSEPERVVPRLDPDINGRRYWRGPDGSGGLIFQTAF 135
Qy 61 QEDFYHLTPDAQFLAPAFSTEHLGVPQLGTGGSDLRRCFYSCDVNAEFDFAAASLC 120
Db 136 QEDFYHLTPDAQFLAPAFSTEHLGVPQLGTGGSDLRRCFYSCDVNAEFDFAAASLC 195
Qy 121 GGLRGAGYGAEEVVISPLPNASAPAAQNSOGAHLRORRGPVGGSDZTSRCGVASGW 180
Db 196 GGLRGAGYGAEEVVISPLPNASAPAAQNSOGAHLRORRGPVGGSDZTSRCGVASGW 255
Qy 181 NPALLRALDPYKPRRAGFGESESRSSRRSGRAKRFVSIPRYVETLVVADESMVKPHGADLEH 240
Db 256 NPALLRALDPYKPRRAGFGESESRSSRRSGRAKRFVSIPRYVETLVVADESMVKPHGADLEH 315
Qy 241 YLLTLTATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTFRNFCAMOKKLN 300
Db 316 YLLTLTATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTFRNFCAMOKKLN 375
Qy 301 KVSDEHPYWDITALET----- 317
Db 376 KVSDEHPYWDITALETTRQPGESPLPSAERRLTICWAIAGEPALSEGVLARPIASAW 435
Qy 318 ----- 317
Db 436 IRRRGAGTQVEAAPELQFVSKTDRRRRRDEGVGGNGPPLLYGRLSLLGLHLRSPFFKGR 495
Qy 318 -----RQDL 321
Db 496 TEGDPRGSGNPKLYKKGQLENNSPGGLRFLAKMIPGSSGKVAATAKNSQLLGIVRQDL 555
Qy 322 CGAITCTDLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHGLHVFNNPHONVKVCEVF 381
Db 556 CGAITCTDLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHGLHVFNNPHONVKVCEVF 615
Qy 382 GKLRANHMSPTLIQIDRANPWSACSAIITDFLDGSGHDCDCLLDQPSKPSLPEDLRGAS 441
Db 616 GKLRANHMSPTLIQIDRANPWSACSAIITDFLDGSGHDCDCLLDQPSKPSLPEDLRGAS 675

Qy 442 YTLSSQCELAFGVSGKPCPYMOYCTKLWCTGKAKGMVCOYTRHPPWADGTSCEGKLCIK 501
Db 676 YTLSSQCELAFGVSGKPCPYMOYCTKLWCTGKAKGMVCOYTRHPPWADGTSCEGKLCIK 735
Qy 502 GACVERHNLNKH-----VDGSAKWDPYGPSCSRCTCGGVQL 538
Db 736 GACVERHNLNKHRLTDIISPKQLLRLPGLHTTKVDGSAKWDPYGPSCSRCTCGGVQL 795
Qy 539 ARROCTNPTPANGKYCEGVYRKYRSCNLEPCPSSASGSKSFREEOCEAFNHNHSTNRLI 598
Db 796 ARROCTNPTPANGKYCEGVYRKYRSCNLEPCPSSASGSKSFREEOCEAFNHNHSTNRLI 855
Qy 599 LAVAWPKYSGVSPRDKCKLICRANGTGYFYVLAPVKVDGTLCSPDSTSVCVQKCKIKAG 659
Db 856 LAVAWPKYSGVSPRDKCKLICRANGTGYFYVLAPVKVDGTLCSPDSTSVCVQKCKIKAG 915
Qy 659 CDGNLGSKKRFDKCGVCGGDKNSCKVTGLFTKPMHGVNFVVAIPAGASSIDIRQGVKG 718
Db 916 CDGNLGSKKRFDKCGVCGGDKNSCKVTGLFTKPMHGVNFVVAIPAGASSIDIRQGVKG 975
Qy 719 LIGDDNYLALKNSQKYLKNGHEFVSAVERDLVWKGSLRLYSGTGTAVESLQASRPILP 778
Db 976 LIGDDNYLALKNSQKYLKNGHEFVSAVERDLVWKGSLRLYSGTGTAVESLQASRPILP 1035
Qy 779 LTVEVLVSGKMTPPRYRYSFYLKPEPRDKSSHPRGPGSVLHNSVLNSLNOVQPDQR 838
Db 1036 LTVEVLVSGKMTPPRYRYSFYLKPEPRDKSSHPRGPGSVLHNSVLNSLNOVQPDQR 1095
Qy 839 PPARWAGSWGPGSCGSLGKRAVDRCGSAGQRTVPACDAHRPVTACGECPTWE 898
Db 1096 PPARWAGSWGPGSCGSLGKRAVDRCGSAGQRTVPACDAHRPVTACGECPTWE 1155
Qy 899 LSWSPCSKSCGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 950
Db 1156 LSWSPCSKSCGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 1207

RESULT 7
US-60-242-679-1059
; Sequence 1059, Application us/60242679
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steven Istvan
; APPLICANT: Spier, Eugene
; APPLICANT: Greenberg, Simon
; APPLICANT: Brandenberger, Ralph
; APPLICANT: Wang, Yu
; APPLICANT: Dubman, Alex
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: US/60-242,679
; FILE REFERENCE: C1000898-PROV
; CURRENT APPLICATION NUMBER: US/60-242,679
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 2265
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1059
; LENGTH: 1252
; TYPE: PRT
; ORGANISM: HUMAN
US-60-242-679-1059

Query Match 97.7%; Score 5041; DB 27; Length 1252;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 202; Gaps 2;

Qy 1 MLLGILTLTAFAGTAGGSEPERVVPRLDPDINGRRYWRGPDGSDGGLIFQTAF 60
Db 101 MLLGILTLTAFAGTAGGSEPERVVPRLDPDINGRRYWRGPDGSDGGLIFQTAF 160
Qy 61 QEDFYHLTPDAQFLAPAFSTEHLGVPQLGTGGSDLRRCFYSCDVNAEFDFAAASLC 120
Db 161 QEDFYHLTPDAQFLAPAFSTEHLGVPQLGTGGSDLRRCFYSCDVNAEFDFAAASLC 220

```
QY 121 GGLRGAFYGAETVLSPLPNASAPAAORNSOGAHLQRRGVPGPSDDPTSCRGVASGW 180
Db 221 GGLRGAFYGAETVLSPLPNASAPAAORNSOGAHLQRRGVPGPSDDPTSCRGVASGW 280
QY 181 NPAILRALDYPKRRAGFSGESRRSRGKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 281 NPAILRALDYPKRRAGFSGESRRSRGKRFVSIPIRYVETLVVADESVMKFGADLEH 340
QY 241 YLLTLATAARLYRHPSILNPINIVVVKVLLRDSDGPKVTGNAALTLENCAWOKKLN 300
Db 341 YLLTLATAARLYRHPSILNPINIVVVKVLLRDSDGPKVTGNAALTLENCAWOKKLN 400
QY 301 KVS DKHPEYWDTAILFTQ----- 319
Db 401 KVS DKHPEYWDTAILFTQPPGESPLPSAERRLTCIWAIGEPALSEGVLARPIASW 460
QY 320 ----- 319
Db 461 IRRRGAGTQVAAPELQFVSKTDRRRREDGVEGNGGPPILLYGRSLSLGLHLLRSPFKGCR 520
QY 320 -----DLCGATTCDTLGMADVGTMCOPKRS 344
Db 521 TEGLSPLSVNRHLELQGEAFHSWSDMHQROQOEVDLDCGATTCDTLGMADVGTMCOPKRS 580
QY 345 CSVIEDDGLPSAFTTAHELGHVFNPHDNVVKCEVEFGKLRANHMMSPTLIQIDRANPWS 404
Db 581 CSVIEDDGLPSAFTTAHELGHVFNPHDNVVKCEVEFGKLRANHMMSPTLIQIDRANPWS 640
QY 405 ACSAAITDPLDSHGDCILLDQSPKIPISLPEDLPGASVYTLSSQCELAFAFGVSGKPCPYMQY 464
Db 641 ACSAAITDPLDSHGDCILLDQSPKIPISLPEDLPGASVYTLSSQCELAFAFGVSGKPCPYMQY 700
QY 465 CTKLWCTGKAKQGVQOTRHFHWADGTSCGEGKCLKACVERHNLNKRH----- 514
Db 701 CTKLWCTGKAKQGVQOTRHFHWADGTSCGEGKCLKACVERHNLNKRHSLMTFLLIS 760
QY 515 -----VDGS 518
Db 761 RRGKPTGSKELYKLIUVNSNIPPLFSAHSAKQALTDIISPKQLLLRLPNGLHTTKVDGS 820
QY 519 WAKWDYPGCSRTCCGGVQOLARROCTNPTPANGKGYCEGVVKYRSCNLEPCPSSASGKS 578
Db 821 WAKWDYPGCSRTCCGGVQOLARROCTNPTPANGKGYCEGVVKYRSCNLEPCPSSASGKS 880
QY 579 FREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCKLICRANGTGYFVLAPKVVDG 638
Db 881 FREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCKLICRANGTGYFVLAPKVVDG 940
QY 639 TLCSPDSTSVQVQCKIKACGDNLGSKKREDKCGVCGDNKSKCKVTGLFTKPMHGYNF 698
Db 941 TLCSPDSTSVQVQCKIKACGDNLGSKKREDKCGVCGDNKSKCKVTGLFTKPMHGYNF 1000
QY 699 VVAIPAGASSIDIRQRYKGLIGDDNYLALNKSQKYLNGHFVVSVAVERDLVYKGLSLR 758
Db 1001 VVAIPAGASSIDIRQRYKGLIGDDNYLALNKSQKYLNGHFVVSVAVERDLVYKGLSLR 1060
QY 759 YSGTGTAVESLQASRPILSPLTVEVLSVCKMTPPRVRYSYFLPKREDKSSHPKOPRGP 818
Db 1061 YSGTGTAVESLQASRPILSPLTVEVLSVCKMTPPRVRYSYFLPKREDKSSHPKOPRGP 1120
QY 819 SVLHNSVLSLSNQVEQDPRPARVAGSWGPCASCSGSGLOKRAVDCRCSAGQRTVPAC 878
Db 1121 SVLHNSVLSLSNQVEQDPRPARVAGSWGPCASCSGSGLOKRAVDCRCSAGQRTVPAC 1180
QY 879 DAAHRPVETOACGEPCTWELSAWSPCSKSGRGFORSLKCVGHGGRLLARQOCNLRHK 938
Db 1181 DAAHRPVETOACGEPCTWELSAWSPCSKSGRGFORSLKCVGHGGRLLARQOCNLRHK 1240
QY 939 POELDFCVLRPC 950
Db 1241 POELDFCVLRPC 1252
```

```
RESULT 8
US-60-207-315-292
; Sequence 292, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 292
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-292

Query Match 97.6%; Score 5038; DB 27; Length 1158;
Best Local similarity 82.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 208; Gaps 2;

QY 1 MLLGLITLTAFAAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGQGLIFOITAF 60
Db 1 MLLGLITLTAFAAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGQGLIFOITAF 60
QY 61 QEDFYHLTPDQAQFLAPAFSTHGLGVPLQGLTGGSSDLRRCFYSGDVNAEPDFAAVSLC 120
Db 61 QEDFYHLTPDQAQFLAPAFSTHGLGVPLQGLTGGSSDLRRCFYSGDVNAEPDFAAVSLC 120
QY 121 GGLRGAFYGAETVLSPLPNASAPAAORNSOGAHLQRRGVPGPSDDPTSCRGVASGW 180
Db 121 GGLRGAFYGAETVLSPLPNASAPAAORNSOGAHLQRRGVPGPSDDPTSCRGVASGW 180
QY 181 NPAILRALDYPKRRAGFSGESRRSRGKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDYPKRRAGFSGESRRSRGKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLATAARLYRHPSILNPINIVVVKVLLRDSDGPKVTGNAALTLENCAWOKKLN 300
Db 241 YLLTLATAARLYRHPSILNPINIVVVKVLLRDSDGPKVTGNAALTLENCAWOKKLN 300
QY 301 KVS DKHPEYWDTAILFT----- 317
Db 301 KVS DKHPEYWDTAILFTQRRSKSPLYRAMPPSLNLVDLFLQLLQCGVSRSPSPFSAERFL 360
QY 318 ----- 317
Db 361 TCITWATGGEPALESGVGLARPIASWIRRRGAGTQVAAPELQFVSKTDRRREDGVEG 420
QY 318 ----- 317
Db 421 GNGPPLLYGLSLGLHLLRSPKCGKRTGSDPCGSKNPKLFYKKGQLENNSPGSLRFLAKM 480
QY 318 -----RDLCGATTCDTLGMADVGTMCDDPKRSCSVIEDDGLS 355
Db 481 IPGSSGKVAATAKNSQLLGIIVRODLGAGATTCDTLGMADVGTMCDDPKRSCSVIEDDGLS 540
QY 356 AFTTAHELGHVFNPHDNVVKCEVEFGKLRANHMMSPTLIQIDRANPWSACSAAITDEL 415
Db 541 AFTTAHELGHVFNPHDNVVKCEVEFGKLRANHMMSPTLIQIDRANPWSACSAAITDEL 600
QY 416 DSHGDCILLDQSPKIPISLPEDLPGASVYTLSSQCELAFAFGVSGKPCPYMQYCTKLWCTGKAK 475
Db 601 DSHGDCILLDQSPKIPISLPEDLPGASVYTLSSQCELAFAFGVSGKPCPYMQYCTKLWCTGKAK 660
QY 476 GQWVCOTRHFHWADGTSCGEGKCLKACVERHNLNKRH----- 514
Db 661 GQWVCOTRHFHWADGTSCGEGKCLKACVERHNLNKRHSLMTFLLIS 720
QY 515 --VDGSKAKWDPYGPCSRTCCGGVQOLARROCTNPTPANGKGYCEGVVKYRSCNLEPCPS 572
Db 515 -----
```


Db 721 TKVDGSAKWDPYPCGSRCTGGGQVLARQCTNPTPANGKVKCEGVVRVKYRSCNLEPCPS 780
QY 573 SASGKSFREOQCAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANGTYFYVILA 632
Db 781 SASGKSFREOQCAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANGTYFYVILA 840
QY 633 PKVVDGTLGSPDSTSCVCGKCKIACDGNLGSKKRFDKGVCGGDNKSKKVTGLFTKP 692
Db 841 PKVVDGTLGSPDSTSCVCGKCKIACDGNLGSKKRFDKGVCGGDNKSKKVTGLFTKP 900
QY 693 MHGYNFVAIPAGASSIDIRQYKGLIGDNYLALKNSSGKYLNGHFVSAVERDLVV 752
Db 901 MHGYNFVAIPAGASSIDIRQYKGLIGDNYLALKNSSGKYLNGHFVSAVERDLVV 960
QY 753 KGSLLRYSGTGTAVERSLQASRPILPTVEVLVSGKWTTPRVYSFLPKPEDEKSSHP 812
Db 961 KGSLLRYSGTGTAVERSLQASRPILPTVEVLVSGKWTTPRVYSFLPKPEDEKSSHP 1020
QY 813 KDPGSPVLSNLSLSNQVEQDDPPARVWAGSWGPCSSGSGLOKAVDCRGSAGQ 872
Db 1021 KDPGSPVLSNLSLSNQVEQDDPPARVWAGSWGPCSSGSGLOKAVDCRGSAGQ 1080
QY 873 RTVPACDAARHPVETQACGPECTWELSAWSPCKSCGGRFQRRSLKCVSHGGRLLARDO 932
Db 1081 RTVPACDAARHPVETQACGPECTWELSAWSPCKSCGGRFQRRSLKCVSHGGRLLARDO 1140
QY 933 CNLHRKQELDFCVLRPC 950
Db 1141 CNLHRKQELDFCVLRPC 1158

RESULT 9

US-10-093-463-28

Sequence 28 Application US/10093463

GENERAL INFORMATION:

APPLICANT: Padigar, Muralidhara
APPLICANT: Shenoy, Suresh
APPLICANT: Kekuda, Ramesh
APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal
APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glenda
APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Zerhusen, Bryan
APPLICANT: Tchernev, Velizar
APPLICANT: Gangolli, Esha
APPLICANT: Vernet, Corine
APPLICANT: Pena, Carol
APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patturajan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids, and Methods of Use.
FILE OF INVENTION: Encoding The Antigens, and Methods of Use.
FILE REFERENCE: 21402-290A (Cura 590AT)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092

PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 28
LENGTH: 924
TYPE: PRT
ORGANISM: Homo sapiens
US-10-093-463-28

Query Match 95.3%; Score 4921; DB 24; Length 924;

Best Local Similarity 95.9%; Pred. No. 0;

Matches 914; Conservative 2; Mismatches 5; Indels 32; Gaps 3;

QY 1 MLLGLITLAFAGTAGGSEPEREVVPIRLDPDINGRYYWRGPDSDQGLIFQITAF 60

Db 1 MLLGLITLAFAGTAGGSEPEREVVPIRLDPDINGRYYWRGPDSDQGLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAVSLC 120

Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAVSLC 120

QY 121 GGLRGAFYRGAEYVISPPLNASAPAAQNSOGAHLQRRGYPGGPSGDPSTFCGVASGW 180

Db 121 GGLRGAFYRGAEYVISPPLNASAPAAQNSOGAHLQRRGYPGGPSGDPSTFCGVASGW 180

QY 181 NPAILRALDPYKPRRAGFGESESRSSRAKRFVSIPIRYVETLVVADESMVKFHGADLEH 240

Db 181 NPAILRALDPYKPRRAGFGESESRSSRAKRFVSIPIRYVETLVVADESMVKFHGADLEH 240

QY 241 YLLTLATAARLYRHPSSIILNPINVVVYVLLLRDRDSGPKVTGNAALTURNCAMQKKLN 300

Db 241 YLLTLATAARLYRHPSSIILNPINVVVYVLLLRDRDSGPKVTGNAALTURNCAMQKKLN 300

QY 301 KVSDEKPEYWDTAILETRQDLGATTCDTILGMADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360

Db 301 KVSDEKPEYWDTAILETRQDLGATTCDTILGMADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360

QY 361 HELGHVFNPHDNVYKVEEVEFGKLRANHMSPPTLIQIDRANPWSACSAIITDFLDSHG 420

Db 361 HELGHVFNPHDNVYKVEEVEFGKLRANHMSPPTLIQIDRANPWSACSAIITDFLDSHG 420

QY 421 DCLLDQPSKPISLPDLPGASYTLSSQCELAFAVGSGSKPCPYMVCYTKLWCTGKRAKGQWVC 480
Db 421 DCLLDQPSKPISLPDLPGASYTLSSQCELAFAVGSGSKPCPYMVCYTKLWCTGKRAKGQWVC 480
QY 481 QTRHFPNADGTSCEGKLCILKGCACVERHNLNKH---RVDGSWAKWDPYGPCSTCGGQVQ 537
Db 481 QTRHFPNADGTSCEGKLCILKGCACVERHNLNKHSSQVDSWAKWDPYGCSTCGGQVQ 540
QY 538 LARROCTNPTPANGKCYCEGVRYKRSCLNEPCPSSASGSKSFREOCEAFNGYHSTNRL 597
Db 541 LARROCTNPTPANGKCYCEGVRYKRSCLNEPCPS--SGKSFREOCEAFNGYHSTNRL 598
QY 598 TLAVAWPKYSGVSPRDKCLICRANGTGYFYVLAKPVVDGTLCSPDSTSVCVQGCICA 657
Db 599 TLAVAWPKYSGVSPRDKCLICRANGTGYFYVLAKPVVDGTLCSPDSTSVCVQGCICA 658
QY 658 GCDNLGSKKRFDKCGVCGDNKCKVKTGLFTKPMHGYNFVVAIPAGASSIDIORGYK 717
Db 659 GCDNLGSKKRFDKCGVCGDNKCKVKTGLLSPARHGYNFVVAIPAGASSIDIORGYK 718
QY 718 GLIGDDNYLALKNSQGYKLLNGHFVSAVERDLVWKGSLRLRYSGTGTAVERSLQASRPIL 777
Db 719 GLIGDDNYLALKNSQGYKLLNGHFVSAVERDLVWKGSLRLRYSGTGTAVERSLQASRPIL 778
QY 778 PLTVEVLVSCKMTPPRVRYSFYLPKEPREDKSSHDPKDPGCPVLSVLSLSNOVEOPDD 837
Db 779 PLTVEVLVSCKMTPPRVRYSFYLPKEPREDKSSH----- 812
QY 838 RPARWVAGSWGPCASCGSLQKRAVDCRGSGAGQRTVPACDAHRPVTQACGEPCTW 897
Db 813 -PRAWVAGSWGPCASCGSLQKRAVDRWGSAGQRTVPACDAHRPVTQACGEPCTW 871
QY 898 ELSAWSPCSKSGRGQRRLSKCVGHGRLLRDQCNLHRKPOELDFCVLRPC 950
Db 872 ELSAWSPCSKSGRGQRRLSKCVGHGRLLRDQCNLHRKPOELDFCVLRPC 924

RESULT 10

US-60-208-020-103
; Sequence 103, Application US/60208020
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000630
; CURRENT APPLICATION NUMBER: US/60/208,020
; CURRENT FILING DATE: 2000-05-31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 884
; TYPE: PRT
; ORGANISM: HUMAN
US-60-208-020-103

Query Match 92.2%; Score 4758; DB 27; Length 884;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 878; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 44 GPDSGDQGLIFQITAFQEDFYHLHTPDQAFLAPAFSTEHLGVPLQGLTGSSDLRRCFY 103
Db 1 GPDSGDQGLIFQITAFQEDFYHLHTPDQAFLAPAFSTEHLGVPLQGLTGSSDLRRCFY 60
QY 104 SGDVNAEPDSFAVSLCGLRGAFYRGAEYVISPPLNAPAAQNSOGAHLQRRGVP 163
Db 61 SGDVNAEPDSFAVSLCGLRGAFYRGAEYVISPPLNAPAAQNSOGAHLQRRGVP 120
QY 164 GPGSDPTSRGCVASGNPAIRALDPTKPRAGFSGESRRSRGRKRFVSIPIRYVETL 223
Db 121 GPGSDPTSRGCVASGNPAIRALDPTKPRAGFSGESRRSRGRKRFVSIPIRYVETL 180
QY 224 VVADESMWKFHGCADLEHYLLTLAARLYRHPSLNINIVVVKVLLLRDRDSKPVIG 283

Db 181 VVADESMWKFHGCADLEHYLLTLAARLYRHPSLNINIVVVKVLLLRDRDSKPVIG 240
QY 284 NAALTLRNFCAWQKLNKVSOKHPEYWDATILFTRQDLCCGATTCDTLGMADVGMCDPKR 343
Db 241 NAALTLRNFCAWQKLNKVSOKHPEYWDATILFTRQDLCCGATTCDTLGMADVGMCDPKR 300
QY 344 SCSTVEDDGLPSAFTTAHELGHVFNPHDNVKNVCEEVFGKLRANHMSPPTLIODRANPW 403
Db 301 SCSTVEDDGLPSAFTTAHELGHVFNPHDNVKNVCEEVFGKLRANHMSPPTLIODRANPW 360
QY 404 SACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASYTLSSQCELAFAVGSGSKPCPYM 463
Db 361 SACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASYTLSSQCELAFAVGSGSKPCPYM 420
QY 464 YCTKLWCTGKAKGQWVCQTRHFPNADGTSCEGKLCILKGCACVERHNLNKH-----VDG 517
Db 421 YCTKLWCTGKAKGQWVCQTRHFPNADGTSCEGKLCILKGCACVERHNLNKHVVSSSQVIG 480
QY 518 SWAKWDPYGPCSRTCGGVQLARQCTNPTPANGKCYCEGVRYKRSCLNEPCPSSASGK 577
Db 481 SWAKWDPYGPCSRTCGGVQLARQCTNPTPANGKCYCEGVRYKRSCLNEPCPSSASGK 540
QY 578 SFREOCEAFNGYHSTNRLTLAVAWPKYSGVSPRDKCLICRANGTGYFYVLAPKVV 637
Db 541 SFREOCEAFNGYHSTNRLTLAVAWPKYSGVSPRDKCLICRANGTGYFYVLAPKVV 600
QY 638 GTLCSPDSTSVCVQGCILKAGCDNLGSKKRFDKCGVCGDNKCKVKTGLFTKPMHGYN 697
Db 601 GTLCSPDSTSVCVQGCILKAGCDNLGSKKRFDKCGVCGDNKCKVKTGLFTKPMHGYN 660
QY 698 FWVAIPAGASSIDIORGYKGLIGDDNYLALKNSQGYKLLNGHFVWSAVERDLVWKGSL 757
Db 661 FWVAIPAGASSIDIORGYKGLIGDDNYLALKNSQGYKLLNGHFVWSAVERDLVWKGSL 720
QY 758 RYSGTGTAVERSLQASRPILFTEVLSVGMKTPPRVRYSFYLPKEPREDKSSHDPKDP 817
Db 721 RYSGTGTAVERSLQASRPILFTEVLSVGMKTPPRVRYSFYLPKEPREDKSSHDPKDP 780
QY 818 PSVLHNSVLSLSNOVEOPDDPPARWVAGSWGPCASCGSLQKRAVDCRGSGAQRTVFA 877
Db 781 PSVLHNSVLSLSNOVEOPDDPPARWVAGSWGPCASCGSLQKRAVDCRGSGAQRTVFA 840
QY 878 CDAHRPVETOACGEPCTWELSAWSPCSKSGRGFQRRSLKCV 921
Db 841 CDAHRPVETOACGEPCTWELSAWSPCSKSGRGFQRRSLKCV 884

RESULT 11

US-60-209-043-133
; Sequence 133, Application US/60209043
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000641
; CURRENT APPLICATION NUMBER: US/60/209,043
; CURRENT FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133
; LENGTH: 884
; TYPE: PRT
; ORGANISM: HUMAN
US-60-209-043-133

Query Match 92.2%; Score 4758; DB 27; Length 884;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 878; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
QY 44 GPDSGDQGLIFQITAFQEDFYHLHTPDQAFLAPAFSTEHLGVPLQGLTGSSDLRRCFY 103

```
Db 1 GPEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFY 60
QY 104 SCDVNAEPDPSAAVSLCGLGAGCYGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVP 163
Db 61 SCDVNAEPDPSAAVSLCGLGAGCYGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVP 120
QY 164 GPGSDPTSRGCVASGWNPAIILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSPRYVETL 223
Db 121 GPGSDPTSRGCVASGWNPAIILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSPRYVETL 180
QY 224 VYADESMVKFHGADLEHYLLTAAARLYRHPISILNPNINIVVVKVLLLRDSDGPKVTG 283
Db 181 VYADESMVKFHGADLEHYLLTAAARLYRHPISILNPNINIVVVKVLLLRDSDGPKVTG 240
QY 284 NAALTLRNFCAWOKLKNVSKDHPYWDTAILETRQDLGCAATCTDGLGMADVGTMCDDPKR 343
Db 241 NAALTLRNFCAWOKLKNVSKDHPYWDTAILETRQDLGCAATCTDGLGMADVGTMCDDPKR 300
QY 344 SCSEVIEDDGLPSAFTTAHELGHVFNMPHDNVKVEEYFGKLRANHMSPITLIQIDRANPW 403
Db 301 SCSEVIEDDGLPSAFTTAHELGHVFNMPHDNVKVEEYFGKLRANHMSPITLIQIDRANPW 360
QY 404 SACSAAITDFLDSHGDCLLDQPSKPIISLPEDLPAGASYTILSQOCELAFVGSKPCPYMQ 463
Db 361 SACSAAITDFLDSHGDCLLDQPSKPIISLPEDLPAGASYTILSQOCELAFVGSKPCPYMQ 420
QY 464 YCTKLWCTGKAKQWQVCTRHFPWADGTSCEGKLCGLKACVERHNLNKHRR-----VDG 517
Db 421 YCTKLWCTGKAKQWQVCTRHFPWADGTSCEGKLCGLKACVERHNLNKHRRVSSSQVDG 480
QY 518 SWAKWDYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 577
Db 481 SWAKWDYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 540
QY 578 SPREQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 637
Db 541 SPREQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 600
QY 638 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDNKSKCKVTGLFTKPMHGYN 697
Db 601 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDNKSKCKVTGLFTKPMHGYN 660
QY 698 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQKYLINGHFVYSAVERDLVVKGSLL 757
Db 661 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQKYLINGHFVYSAVERDLVVKGSLL 720
QY 758 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPRVRYSFYLPEKPEDEKSSHPKDPRG 817
Db 721 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPRVRYSFYLPEKPEDEKSSHPKDPRG 780
QY 818 PSVLHNSVLSNVOEQDPRPARWAGSWGPCSASGSLQKRAVDCRSAGQRTVPA 877
Db 781 PSVLHNSVLSNVOEQDPRPARWAGSWGPCSASGSLQKRAVDCRSAGQRTVPA 840
QY 878 CDAHRPVETQACGPECPWTWELSAWSPCSKSGRGRFQRRSLKCV 921
Db 841 CDAHRPVETQACGPECPWTWELSAWSPCSKSGRGRFQRRSLKCV 884
```

```
RESULT 12
US-10-163-316-2
; Sequence 2, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-316-2

Query Match      82.3%; Score 4248.5; DB 25; Length 823;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 792; Conservative 0; Mismatches 1; Indels 23; Gaps 1;

QY 1 MLLGLGILTLAFAGRTAGGSEPEREVVVPIRLDDINDINGRRYYWRGPDSDGGLIFQITAF 60
Db 1 MLLGLGILTLAFAGRTAGGSEPEREVVVPIRLDDINDINGRRYYWRGPDSDGGLIFQITAF 60
QY 61 QEDFYHLHLPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDPSFAAVSLC 120
Db 61 QEDFYHLHLPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDPSFAAVSLC 120
QY 121 GGLRGAFGYRGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVPGPGSDPTSRGCVASGW 180
Db 121 GGLRGAFGYRGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVPGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSPRYVETILVVADESMVKFHGADLEH 240
Db 181 NPAILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSPRYVETILVVADESMVKFHGADLEH 240
QY 241 YLLTLATAARLYRHPISILNPNINIVVVKVLLLRDSDGPKVTGNAALTILRNFCAWOKKLN 300
Db 241 YLLTLATAARLYRHPISILNPNINIVVVKVLLLRDSDGPKVTGNAALTILRNFCAWOKKLN 300
QY 301 KVS DKHPYWDTAILETRQDLGCAATCTDGLGMADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360
Db 301 KVS DKHPYWDTAILETRQDLGCAATCTDGLGMADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNMPHDNVKVEEYFGKLRANHMSPITLIQIDRANPWSACSAALITDFLDSHG 420
Db 361 HELGHVFNMPHDNVKVEEYFGKLRANHMSPITLIQIDRANPWSACSAALITDFLDSHG 420
QY 421 DCLLDQPSKPIISLPEDLPAGASYTILSQOCELAFVGSKPCPYMOYCTKLWCTGKAKQWVC 480
Db 421 DCLLDQPSKPIISLPEDLPAGASYTILSQOCELAFVGSKPCPYMOYCTKLWCTGKAKQWVC 480
QY 481 QTRHFPWADGTSCEGKLCGLKACVERHNLNKHRR-----VDG 517
Db 481 QTRHFPWADGTSCEGKLCGLKACVERHNLNKHRRPTDIISPKQLLLPLNGLHTTQVDG 540
QY 518 SWAKWDYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 577
Db 541 SWAKWDYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 600
QY 578 SFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 637
Db 601 SFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 660
QY 638 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDNKSKCKVTGLFTKPMHGYN 697
Db 661 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDNKSKCKVTGLFTKPMHGYN 720
QY 698 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQKYLINGHFVYSAVERDLVVKGSLL 757
Db 721 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQKYLINGHFVYSAVERDLVVKGSLL 780
QY 758 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPR 793
Db 781 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPR 816

RESULT 13
US-60-297-863-2
; Sequence 2, Application US/60297863
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses
```

; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI2001-025P1(M)
; CURRENT APPLICATION NUMBER: US/60/297,863
; CURRENT FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-297-863-2

Query Match 82.3%; Score 4248.5; DB 27; Length 823;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 792; Conservative 0; Mismatches 1; Indels 23; Gaps 1;
QY 1 MLLGLILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60
Db 1 MLLGLILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60
QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSDLRRCFYSGDVNAEPDPSFAAVSLC 120
Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSDLRRCFYSGDVNAEPDPSFAAVSLC 120
QY 121 GGLRGAFYRGAEVYISPLPNASAPAAQRNSOGAHLLORRGYPGPGSDPTSRGCVASGW 180
Db 121 GGLRGAFYRGAEVYISPLPNASAPAAQRNSOGAHLLORRGYPGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGFGESRRSRSSRAKRFVSIPIRYVETLVVADESNNVAFHGADLEH 240
Db 181 NPAILRALDPYKPRRAGFGESRRSRSSRAKRFVSIPIRYVETLVVADESNNVAFHGADLEH 240
QY 241 YLTLTLATAARLYRHPISILNPINIVVVKVLLLRDSDGSKPVGTGNAALTIRNFCAMOKKLN 300
Db 241 YLTLTLATAARLYRHPISILNPINIVVVKVLLLRDSDGSKPVGTGNAALTIRNFCAMOKKLN 300
QY 301 KVSQKHPYWDYTAILTRQDLGCATTCDTLGMADVGTCDPKRSCSVIEDDGLPSAFTTA 360
Db 301 KVSQKHPYWDYTAILTRQDLGCATTCDTLGMADVGTCDPKRSCSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNPHDNVNVKCEEVFGKLRANHMSPITLIQIDRANPWSACSAAIITDFLDSGHG 420
Db 361 HELGHVFNPHDNVNVKCEEVFGKLRANHMSPITLIQIDRANPWSACSAAIITDFLDSGHG 420
QY 421 DCLLDQSPKIPISLPEDLPGASYTLSSOCELAFGVSKPCPYQWYCTKLWCTKAKGQWVC 480
Db 421 DCLLDQSPKIPISLPEDLPGASYTLSSOCELAFGVSKPCPYQWYCTKLWCTKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKHRR-----VDG 517
Db 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKHRRPTDILSPKQLLLRLPGLHTTQVDG 540
QY 518 SWAKWDPYGPCSTCGGQVQLARRQCTNTPANGKYGCVRVKYSRCHLEPCPSASGK 577
Db 541 SWAKWDPYGPCSTCGGQVQLARRQCTNTPANGKYGCVRVKYSRCHLEPCPSASGK 600
QY 578 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCLICRANGTGYFYVLAPKVD 637
Db 601 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCLICRANGTGYFYVLAPKVD 660
QY 638 GTLCSPDSTSVCVQKRCIKAGCDNGLGSKKRFCDKVCVCGDKNKCKVTGFTKPMHGVN 697
Db 661 GTLCSPDSTSVCVQKRCIKAGCDNGLGSKKRFCDKVCVCGDKNKCKVTGFTKPMHGVN 720
QY 698 FVVAIPAGASSIDIRORYKGLIGDDNYLALNOSCKYLLNGHFVVSAVERDLVVKGSLL 757
Db 721 FVVAIPAGASSIDIRORYKGLIGDDNYLALNOSCKYLLNGHFVVSAVERDLVVKGSLL 780
QY 758 RYSGTGTAVESIQASRPILPLTVEVLSVGMTPPR 793
Db 781 RYSGTGTAVESIQASRPILPLTVEVLSVGMTPPR 816

RESULT 14
US-10-093-463-30
; Sequence 30, Application US/10093463
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennnda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zernusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03

;; PRIOR APPLICATION NUMBER: 60/275,578
;; PRIOR FILING DATE: 2001-03-13
;; NUMBER OF SEQ ID NOS: 370
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 30
;; LENGTH: 755
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-093-463-30

Query Match 76.88; Score 3963.5; DB 24; Length 755;
Best Local Similarity 79.58; Pred. No. 0;
Matches 755; Conservative 0; Mismatches 0; Indels 195; Gaps 1;

QY 1 MLLGLTLAFAGTAGGSEPERVVPVIRLDPDINGRRYWRGPDSDGGLFIQITAF 60
DB 1 MLLGLTLAFAGTAGGSEPERVVPVIRLDPDINGRRYWRGPDSDGGLFIQITAF 60

QY 61 QEDFYLHLPDPAQFLAPAFSTHGLGVPLQGLTGSSDLRCKFCYSGDVNAEPDFAVSLC 120
DB 61 QEDFYLHLPDPAQFLAPAFSTHGLGVPLQGLTGSSDLRCKFCYSGDVNAEPDFAVSLC 120

QY 121 GGLGAFGYRGAEEYISPLPNASAPAAQRNSQGAHLQRRGVPGPSGDPTRSGCVASGW 180
DB 121 GGLGAFGYRGAEEYISPLPNASAPAAQRNSQGAHLQRRGVPGPSGDPTRSGCVASGW 180

QY 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKRFFVSPRYVETLVADSMVKEFHGADLEH 240
DB 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKRFFVSPRYVETLVADSMVKEFHGADLEH 240

QY 241 YLLTLATAARLYRHPSILNPINIVVVKVLLLRDRSGPKVTGNAALTLRNCAMOKKLN 300
DB 241 YLLTLATAARLYRHPSILNPINIVVVKVLLLRDRSGPKVTGNAALTLRNCAMOKKLN 300

QY 301 KVSDDHPEYWDTAILFTQDLGCGATTCTDGLMADVGTMCDPKRSCEVIEDGLPSAFTTA 360
DB 301 KVSDDHPEYWDTAILFTQDLGCGATTCTDGLMADVGTMCDPKRSCEVIEDGLPSAFTTA 360

QY 361 HELGHVFNMPHDNVKVCVEVFGKLRANHMMSPTLIQIDRANPNWSACSAAIITDFDSHG 420
DB 320 ----- 319

QY 421 DCLLDQSPKPSLPELDPGASYTLISQOCELAFGVGSKPCPYMQYCTKLWCTGKAKQWVC 480
DB 320 ----- 319

QY 481 QTRHFPWADGTCGEGKLCCLKGACVERHNLNKHVRDGSMAKWDYPGCSRTGGGVQLAR 540
DB 320 ----- 319

QY 541 RQCTNPTPANGKCYCEGYRVYKRSCLNLEPCPSASGSKSFREOCEAFNGYHNSTNKLTLA 600
DB 346 RQCTNPTPANGKCYCEGYRVYKRSCLNLEPCPSASGSKSFREOCEAFNGYHNSTNKLTLA 405

QY 601 VAWPKYSGVSPDRCKLICRANGTGYFYVLA PKVVDGTLCSPDSTSVQCKICAKGD 660
DB 406 VAWPKYSGVSPDRCKLICRANGTGYFYVLA PKVVDGTLCSPDSTSVQCKICAKGD 465

QY 661 GNLGSKKRPDKCGVCGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
DB 466 GNLGSKKRPDKCGVCGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 525

QY 721 GDDNYLALKNISOGKYLKNGHPVSAVERDLVVGSLRLYSGTGTAVESLCASRPILPLT 780
DB 526 GDDNYLALKNISOGKYLKNGHPVSAVERDLVVGSLRLYSGTGTAVESLCASRPILPLT 585

QY 781 VEVLVSGKMTPPRVRYSTYLPKPEDEKSSHPKDPGRGPSVLHNSVLSLNSQVQPDORPP 840
DB 586 VEVLVSGKMTPPRVRYSTYLPKPEDEKSSHPKDPGRGPSVLHNSVLSLNSQVQPDORPP 645

QY 841 ARWAGSWGPCASGSGLOKRAYDVCGRSAGQRTVPACDAHRPVTQACGECPTWELS 900
DB 646 ARWAGSWGPCASGSGLOKRAYDVCGRSAGQRTVPACDAHRPVTQACGECPTWELS 705

QY 901 AWSFPCSKSGRGFORRSCLKVGHGRLRLARDQCNLHRKPOELDFCVLRPC 950
DB 706 AWSFPCSKSGRGFORRSCLKVGHGRLRLARDQCNLHRKPOELDFCVLRPC 755

RESULT 15
US-60-206-028-172
; Sequence 172, Application US/60206028
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000571
; CURRENT APPLICATION NUMBER: US/60/206,028
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 472
; TYPE: PRT
; ORGANISM: HUMAN
US-60-206-028-172

Query Match 49.04; Score 2528; DB 27; Length 472;
Best Local Similarity 99.44; Pred. No. 5.5e-200;
Matches 469; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 48 SGDOGLIFQITAFQEDFYLHLPDPAQFLAPAFSTHGLGVPLQGLTGSSDLRCKFCYSGEV 107
DB 1 SGDOGLIFQITAFQEDFYLHLPDPAQFLAPAFSTHGLGVPLQGLTGSSDLRCKFCYSGEV 60

QY 108 NAEPPDFAAVSLCGGLRGAFYRGAEEYISPLPNASAPAAQRNSQGAHLQRRGVPGPS 167
DB 61 NAEPPDFAAVSLCGGLRGAFYRGAEEYISPLPNASAPAAQRNSQGAHLQRRGVPGPS 120

QY 168 GDPTSCRGVASGWNPAILRALDPYKPRRAGGSESRSSRRSGRAKRFFVSPRYVETLVAD 227
DB 121 GDPTSCRGVASGWNPAILRALDPYKPRRAGGSESRSSRRSGRAKRFFVSPRYVETLVAD 180

QY 228 ESMVKFEGADLEHLLTLATAARLYRHPSILNPINIVVVKVLLLRDRSGPKVTGNAAL 287
DB 181 ESMVKFEGADLEHLLTLATAARLYRHPSILNPINIVVVKVLLLRDRSGPKVTGNAAL 240

QY 288 TLRNCAWQKLNKYSKDHPEYWDTAILFTQDLGCGATTCTDGLMADVGTMCDPKRSCEV 347
DB 241 TLRNCAWQKLNKYSKDHPEYWDTAILFTQDLGCGATTCTDGLMADVGTMCDPKRSCEV 300

QY 348 IEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEVFGKLRANHMMSPTLIQIDRANPNWSACS 407
DB 301 IEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEVFGKLRANHMMSPTLIQIDRANPNWSACS 360

QY 408 AAIITDFLDSGHGDCILLDQPSKPSILPELDPGASYTLISQOCELAFGVGSKPCPYMQYCTK 467
DB 361 AAIITDFLDSGHGDCILLDQPSKPSILPELDPGASYTLISQOCELAFGVGSKPCPYMQYCTK 420

QY 468 LWCTGAKQWVCQTRHFPWADGTCGEGKLCCLKGACVERHNLNKHVRDGSW 519
DB 421 LWCTGAKQWVCQTRHFPWADGTCGEGKLCCLKGACVERHNLNKHVRDGSW 472

Search completed: May 9, 2003, 15:27:31
Job time : 167 secs

Inter Haven Co

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:22:46 ; Search time 97 Seconds
(without alignments)
1674.819 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLIGILTLFAGRTAGGSE.....DQNLHRKQQLDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 787003 seqs, 171007862 residues

Total number of hits satisfying chosen parameters: 787003

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	5162	100.0	950	6	US-10-391-364-77	Sequence 77, Appl
2	5162	100.0	950	7	US-60-453-135-14971	Sequence 14971, A
3	5162	100.0	950	7	US-60-453-050-14971	Sequence 14971, A
4	5117	99.1	952	6	US-10-311-035-11	Sequence 11, Appl
5	2486.5	48.2	967	7	US-10-369-779-18	Sequence 18, Appl
6	2486.5	48.2	967	7	US-60-452-680-23653	Sequence 23653, A
7	2486.5	48.2	967	7	US-60-453-135-14843	Sequence 14843, A
8	2486.5	48.2	967	7	US-60-453-050-14843	Sequence 14843, A
9	2486.5	48.2	967	7	US-60-455-444-8026	Sequence 8026, Ap
10	2486.5	48.2	999	6	US-10-369-779-31	Sequence 31, Appl
11	2485.5	48.1	951	6	US-10-369-779-29	Sequence 29, Appl
12	2485.5	48.1	951	6	US-10-381-793-3	Sequence 3, Appl
13	2485.5	48.1	968	6	US-10-369-779-19	Sequence 19, Appl
14	2485.5	48.1	968	6	US-10-391-364-82	Sequence 82, Appl
15	2482.5	48.1	967	7	US-60-440-068-134	Sequence 134, App
16	2474	47.9	967	6	US-10-369-779-20	Sequence 20, Appl
17	2457.5	47.6	951	6	US-10-369-779-28	Sequence 28, Appl
18	2138.5	41.4	890	5	US-09-949-002-394	Sequence 394, App
19	2138.5	41.4	890	6	US-10-369-779-27	Sequence 27, Appl
20	2135.5	41.4	821	5	US-09-724-676-80844	Sequence 80844, A
21	2135.5	41.4	821	5	US-09-724-676A-80844	Sequence 80844, A
22	2135.5	41.4	887	5	US-09-949-002-534	Sequence 534, App
23	2135.5	41.4	1004	7	US-60-455-444-8141	Sequence 8141, Ap
24	2130.5	41.3	890	5	US-09-724-676-67251	Sequence 67251, A
25	2130.5	41.3	890	5	US-09-724-676A-67251	Sequence 67251, A
26	1965	38.1	1916	6	US-10-274-639-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-391-364-77
; Sequence 77, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; FILE REFERENCE: MPI03-0190NMN
; CURRENT APPLICATION NUMBER: US/10/391,364
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-364-77

Query Match 100.0%; Score 5162; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLLGLILTLAFAGRTAGGSEPERVVVPIRLDPDINGRRYYWRGPEDSGCLQIFQITAF	60
Db	1	MLLGLILTLAFAGRTAGGSEPERVVVPIRLDPDINGRRYYWRGPEDSGCLQIFQITAF	60
Qy	61	QEDFYLHLTPDAQFLAPASTEHLGVPLQGLTGGSSDLRCHYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYLHLTPDAQFLAPASTEHLGVPLQGLTGGSSDLRCHYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFYGAEYVISPINASAPAAQRNSOGAHLQRRGYPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFYGAEYVISPINASAPAAQRNSOGAHLQRRGYPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFESRRRSRSGRAKRFVSIPIRVETILVWDESWKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFESRRRSRSGRAKRFVSIPIRVETILVWDESWKFHGADLEH	240
Qy	241	YLLTLTLATAARLYRHPESILNPINIVVVKVLLLRDRDSGPKVTGNAALTURNFCAWQKKLN	300
Db	241	YLLTLTLATAARLYRHPESILNPINIVVVKVLLLRDRDSGPKVTGNAALTURNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILTFRODLGCATTCTDGLMADVTGMDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILTFRODLGCATTCTDGLMADVTGMDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVEEVFGKLRANHMSPTLQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVEEVFGKLRANHMSPTLQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQOCELAFGVGSKPCPYWOYCTKLWCTCKAKGQWVC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQOCELAFGVGSKPCPYWOYCTKLWCTCKAKGQWVC	480
Qy	481	QTRHFPWADGTSCEGKCLKLGACVERHNLNKHRVDGWSWAKMDPYGPCSRCTCGGVQQLAR	540
Db	481	QTRHFPWADGTSCEGKCLKLGACVERHNLNKHRVDGWSWAKMDPYGPCSRCTCGGVQQLAR	540
Qy	541	ROCTNPTPANGKYCEGVRVYKYSBCNLPCPSSASGKSPREOCFAFNYNHSTNRLTLA	600
Db	541	ROCTNPTPANGKYCEGVRVYKYSBCNLPCPSSASGKSPREOCFAFNYNHSTNRLTLA	600
Qy	601	VAWPKYSGVSPRDKCLICRANGTYFYVLAPKVYDGTLCSPDSTSVCVQKCIKAGCD	660
Db	601	VAWPKYSGVSPRDKCLICRANGTYFYVLAPKVYDGTLCSPDSTSVCVQKCIKAGCD	660
Qy	661	GNLGSKKRFDKCGCGDNKSKKVTGLFTKPMHGNFYVVAIPAGASSITDIQRGYKGLI	720
Db	661	GNLGSKKRFDKCGCGDNKSKKVTGLFTKPMHGNFYVVAIPAGASSITDIQRGYKGLI	720
Qy	721	GDDNYLALKNISOGKYLINGHFVVSAYERDLVKGSLRLRYSGTGAVESLQASRPILPLT	780
Db	721	GDDNYLALKNISOGKYLINGHFVVSAYERDLVKGSLRLRYSGTGAVESLQASRPILPLT	780
Qy	781	VEVLSVGKMTPPRVRSFYFLPKPEPREDKSSHPKDPRGPSVLINSVLVLSNQVEQDDRRP	840
Db	781	VEVLSVGKMTPPRVRSFYFLPKPEPREDKSSHPKDPRGPSVLINSVLVLSNQVEQDDRRP	840
Qy	841	ARWVAGSWGPCSCSGGLQKRAVDCRGAGORTVPACDAAHRPVBTOACGPPCTWELLS	900
Db	841	ARWVAGSWGPCSCSGGLQKRAVDCRGAGORTVPACDAAHRPVBTOACGPPCTWELLS	900
Qy	901	ANWSPKSCSGRGFORLSLKCVCVGHGRLRLARDQCNLHRKFOELDFCVLRPC	950
Db	901	ANWSPKSCSGRGFORLSLKCVCVGHGRLRLARDQCNLHRKFOELDFCVLRPC	950

RESULT 2

RESULT 2
US-60-453-135-14971

US-60-433-133-14971
: Sequence 14971, Application US/60453135

: sequence 14971, Appl:
: GENERAL INFORMATION:

;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele

;; APPLICANT: CARGILL, MICHELE
: APPLICANT: IAKOUBOVA, OLGA

APPLICANT: IAKOUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

Matches	947;	Conservative	0;	Mismatches	3;	Indels	2;	Gaps	2;
Qy	1	MLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF	60						
Db	1	MLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF	60						
Qy	61	QEDFYLHLTPDQAQFLAPAFSTEHLGVLPLOGLTGGSSDLRCFCYSGDVNAEPDPSFAVSLIC	120						
Db	61	QEDFYLHLTPDQAQFLAPAFSTEHLGVLPLOGLTGGSSDLRCFCYSGDVNAEPDPSFAVSLIC	120						
Qy	121	GGURGAFGYRGAEEYVLSPLPNASAPAAQNSOGAHLQRGVGGPGSDPTSRCGVASGW	180						
Db	121	GGURGAFGYRGAEEYVLSPLPNASAPAAQNSOGAHLQRGVGGPGSDPTSRCGVASGW	180						
Qy	181	NPAILRALDPYKPRRAGGEGESRRRRSGRAKRFYSIPRYVETILWADESVMKFGHADLEH	240						
Db	181	NPAILRALDPYKPRRAGGEGESRRRRSGRAKRFYSIPRYVETILWADESVMKFGHADLEH	240						
Qy	241	YLTLTLTAARLYRHPSILNPINIVVKVLLLRDSDGPKVITGNAALTILNFCAMQKKLN	300						
Db	241	YLTLTLTAARLYRHPSILNPINIVVKVLLLRDSDGPKVITGNAALTILNFCAMQKKLN	300						
Qy	301	KVSDKHPEYWDTAILFTTRQDLCCGATCTDGLMADVTGMDCPKSCSVIEDDGLPSAFTTA	360						
Db	301	KVSDKHPEYWDTAILFTTRQDLCCGATCTDGLMADVTGMDCPKSCSVIEDDGLPSAFTTA	360						
Qy	361	HELGHVFNPHDNVNYCEEVFKGLRANHMSPTLIQIDRANPWSACSAAITDFLDSHG	420						
Db	361	HELGHVFNPHDNVNYCEEVFKGLRANHMSPTLIQIDRANPWSACSAAITDFLDSHG	420						
Qy	421	DCLLDQSPKPSILPEDLPGASVITLSQCELAFGVGSKPCPYMOCYTKLACTGKAGQMYC	480						
Db	421	DCLLDQSPKPSILPEDLPGASVITLSQCELAFGVGSKPCPYMOCYTKLACTGKAGQMYC	480						
Qy	481	QTRHFPWADGTCGEGKCLKGACVERINLKNHRVDGWSWAKWDPYGPSCSTCGGVQOLAR	540						
Db	481	QTRHFPWADGTCGEGKCLKGACVERINLKNHRVDGWSWAKWDPYGPSCSTCGGVQOLAR	540						
Qy	541	RQCTNPTPANGKYGCEGVRYKYSRNCLEPCPSSASGSKSFREEOCEAFNGYHNSTNRLTLA	600						
Db	541	RQCTNPTPANGKYGCEGVRYKYSRNCLEPCPSSASGSKSFREEOCEAFNGYHNSTNRLTLA	600						
Qy	601	VAVVPKYSGVSPDKCKLICRANGTGYEYVLAPK - VVDGTLCSPDSTISVCVOGKCLKAGC	659						
Db	601	VAVVPKYSGVSPDKCKLICRANGTGYEYVLAPKVVVDGTLCSPDSTISVCVOGKCLKAGC	660						
Qy	660	DGNLGSKKRFDKCGVCGGDNKCKKVTGLTKPMHGYNFVVAIPAGASSIDIRQRYKGL	719						
Db	661	DGNLGSKKRFDKCGVCGGDNKCKKVTGLTKPMHGYNFVVAIPAGASSIDIRQRYKGL	720						
Qy	720	IGDDNYLALKNSOGKYLLNGHFVWSAVERDLVWKSLLRLYSGTGTAVESLQASRPILLEPL	779						
Db	721	IGDDNYLALKNSOGKYLLNGHFVWSAVERDLVWKSLLRLYSGTGTAVESLQASRPILLEPL	780						
Qy	780	TVEVLSVGKMTTPRVRYSYLPKEPREDKSSHPKDPK - GFSVLHNSVLISLSNOVEQPDOR	838						
Db	781	TVEVLSVGKMTTPRVRYSYLPKEPREDKSSHPKDPKPHRGGPSVLHNSVLISLSNOVEQPDOR	840						
Qy	839	PPARWAGSWGPCASCGSGLQKRAVDCRGSAGORTYPACDAAHRPVETQACGECPTWE	898						
Db	841	PPARWAGSWGPCASCGSGLQKRAVDNRGSGORTYPACDAAHRPVETQACGECPTWE	900						
Qy	899	LSAWSPCSKSCGRGFORRLSKCVGHGGLLARDQCNLHRKPOEDFCVCLRPC	950						
Db	901	LSAWSPCSKSCGRGFORRLSKCVGHGGLLARDQCNLHRKPOEDFCVCLRPC	952						

RESULT 5
US-10-369-779-18
; Sequence 18, Application US/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Comp
; APPLICANT: Sharma, Rahul

```

: APPLICANT: Ramanathan, Chandra
: APPLICANT: Westphal, Ryan
: APPLICANT: Feder, John
: TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
: TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
: TITLE OF INVENTION: ADAM-TS-S1
: FILE REFERENCE: D0204 NP
: CURRENT APPLICATION NUMBER: US/10/369,779
: CURRENT FILING DATE: 2003-02-19
: PRIOR APPLICATION NUMBER: 60/358,151
: PRIOR FILING DATE: 2002-02-19
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 18
: LENGTH: 967
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-369-779-18

```

Query Match	48.2%;	Score 2486.5;	DB 6;	Length 967;
Best Local Similarity	49.0%;	Pred. No. 2.3e-174;		
Matches 486:	Conservative 154;	Mismatches 250;	Indels 101;	Gaps 24;

[illegible]

Db 726 KTSQVTSKPGYHDIITPTGATNIEVKQNRGRNNGSFLAIIKAADGYIILNGDYTL 785
QY 744 SAVERDLVVKSLRYSGTAVESLOASRPILPTVEVLSVGKMTPPKRVYGFYLPKE 803
Db 786 STLEQDIIKGVVLYSGSSALERISFSLKEPLTIQVLTGVALRPKIKYFYVKK 845
QY 804 PREDKSSHKDPGRPSVILHNSVLSNOVEQDPRPPARWAGSWGPCASCGSLQKRA 863
Db 846 ----KES-----FNAIPTS-----AWIEEWGECSSCELGWQRL 878
QY 864 VDCRSAGQRTVPACDAH--RPVETOACGE-PCPTWELSAWSPCSKSGHGFORSILK 920
Db 879 VECDINGO---PASECAKEVPKASTPCADHPCPQWLGWSSCSKTCGKGYKRSILK 935
QY 921 VGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966
RESULT 6
US-60-452-680-23653
; Sequence 23653, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23653
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-23653

Query Match 48.2%; Score 2486.5; DB 7; Length 967;
Best Local Similarity 48.5%; Pred. No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;
QY 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSGQGLIFQITAF 60
Db 36 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTLRLLAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSDLRRCFYSGDVNAEPDSF 114
Db 82 DQQLDLRLPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAGYRGAEYVISPLPNAS---APAAQRNSOGA---HLLQRRGVPGGPS 167
Db 137 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAPGKPPAPLQFHLRR-----NRQ 191
QY 168 GDPTSCRGV-----ASGNPAILRALDPYKPRRAGFGESR 202
Db 192 GDVGTCGVVDDEPRPTKAEFDEDEGEDEGAQWS-----PQDFALQGVGP- 242
QY 203 SRRRSGRAKRFVSPRYVETLVVADESVMVKGADLEHYLTTLATAARLYRHPSILNPI 262
Db 243 TGTGIRKKRFVSSHRYVETMLVADQSMAEFHGSGIKHYLLTLFSVAARLYKHPISIRNSV 302
QY 263 NTIVVKVLLLRDRDSKPVGTGNAALTNRFCAWOKKLNKVSQKHPEYWDYTAILETRDLC 322
Db 303 SLVVVYKILVHDEQKGPEVTSNAULTLRNFCNQKHPPSDRAEHYDIAILLTRDLC 362
QY 323 GATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFETTAHELGHVFMNPHDNVVKYCEEVFG 382
Db 363 GSQTCDTLGMADVGVCDPRSCSVIEDDGLQAETTAHELGHVFMNPHDDAKQOCASLNG 422
QY 383 KLRANHMSPTLIQIDRANPWGASCAAITITFDLSHGDCDLQDQSKPISIPEDLPGASY 442
Db 423 VNQDSHMAISMLNLDHSPQSPSCAYMTTSPLDNGHGECLMDKPNQPIQLPGDLPGTSY 482

QY 443 TLSQOCELAFGVGSKPCP-YMOYCTKLWCTGKAKGQMVQOTRHFPWADGTSCGEGKLCLK 501
Db 483 DANRCQFTGEDSKDPCDAASTCTLWCTGTSGVLVQCTKHFHPWADGTSCGEGKWCIN 542
QY 502 GACVERHNLNKH---RVDSWAKWDPIYGPCSRTCCGGVQLARRQCTNPPTPANGKCYCEGV 558
Db 543 GKCVNKTID-RKHFDTPFHGSGWGMGPWDCSRTCCGGVQYTMRECDNPVPKNGKCYCEG 601
QY 559 RVKYRSCNLBPCSSASGSKFREEOCAFNGYNHSTNRLTLAVAVVPKYSGVSPDRCKL 618
Db 602 RVIRSCNLEDCPDN-NGKTFREEQEAHNEFSKASFGSGPAVEWIPKTAGVSPKDRCKL 660
QY 619 ICRANGTYFYVLAPKVVVDGTLCSPDSTSVYCVQCKIKAGCCDNLGSKKRDKCGVCGG 678
Db 661 ICQAKGIGYFFVLQPKVVDGTGTPCSPDSTSVYCVQCKVAKGCDRIIDSKKKDKCGVCGG 720
QY 679 NKSCKVTGLFTKPMHGYNFVVAIPAGASSIDTRQRYGKGLIGDDNYLALKNQSQKLYLN 738
Db 721 GSTCKKISGVTSAKPGYHDIITPTGATNIEVKQNRGRNNGSFLAIIKAADGYIILN 780
QY 739 GHFVSAVERDLVVKSLRYSGTAVESLOASRPILPTVEVLSVGKMTPPRVYRSP 798
Db 781 GDTLSTLEQDIIKGVVLYSGSSALERISFSLKEPLTIQVLTGVALRPKIKYFY 840
QY 799 YLPKEPREDKSSHPPKDPGRGPSVILHNSVLSNOVEQDPRPPARWAGSWGPCASCGSG 858
Db 841 FVKKK-----KES-----FNAIPTS-----AWIEEWGECSSCELG 873
QY 859 LQKRAVDCRSAGQRTVPACDAH--RPVETOACGE-PCPTWELSAWSPCSKSGRGFOR 915
Db 874 WQRLVRECDINGO---PASECAKEVPKASTPCADHPCPQWLGWSSCSKTCGKGYK 930
QY 916 RSLKCVCHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 931 RSLKCLSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966
RESULT 7
US-60-453-135-14843
; Sequence 14843, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14843
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-14843

Query Match 48.2%; Score 2486.5; DB 7; Length 967;
Best Local Similarity 48.5%; Pred. No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;
QY 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSGQGLIFQITAF 60
Db 36 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTLRLLAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSDLRRCFYSGDVNAEPDSF 114
Db 82 DQQLDLRLPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAGYRGAEYVISPLPNAS---APAAQRNSOGA---HLLQRRGVPGGPS 167
Db 137 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAPGKPPAPLQFHLRR-----NRQ 191

QY 168 GDTSRGCV-----ASGNPAILRALDPYKPRRAGFESR 202
Db 192 GVGCGTGVVDDPRPTGKAETEDEGTEDEGAQWS-----PODPALQGVGP- 242
QY 203 SRRSGRAKRFVSIPIRYVETLVVADESVMKFGADLEHLLTLLATAARLYRHPSILNPI 262
Db 243 TGTGSIIRKRRFVSSHRYVETMLVADOSMAEFHSGGLKHYLLTLFVSAARLYKHPSIRNSV 302
QY 263 NIIVVVKLLLRDRSDGPKVTGNAALTLRNFCAWQKLNKVSQKHPEYWDPTAILFTRODLC 322
Db 303 SLVVVKILVHDQKPEVTSNAALTLRNFCAWQKLNKVSQKHPEYWDPTAILFTRODLC 362
QY 323 GATTCDTLGMADVTCMDPKRSCSVIEDDGLSAFTTAHELGHVFNMPHNVKVEEVEFG 382
Db 363 GSOTCDTLGMADVTCMDPKRSCSVIEDDGLSAFTTAHELGHVFNMPHNVKVEEVEFG 422
QY 383 KLRANHMWSPTLQIDRANPWSACSAIITDFLDGSHGDCLLDOPSKPISLPEDLPASV 442
Db 423 VNQDSHMMASMLNLDHSQWSPSCSAIITDFLDGSHGDCLLDOPSKPISLPEDLPASV 482
QY 443 TLSQCELAFGVSGKPCP-YMQYCTKLWCTGKAGOMVCQTRHFPWADGTSCEGKCLK 501
Db 483 DANROCOFTFGEDSKHCPDAASTCTLWCTGSGVLVQCTKHFPWADGTSCEGKWCIN 542
QY 502 GACVERHNLKH---RVDGSKAKWDYPGCSRTCCGGVQLARRQCTNPTPANGKVCYEGV 558
Db 543 GKCVNKTD-RKHEDTFFHSGWMMGPMWDCSRTCCGGVQYTMRECDNPVFNKGKCYEGK 601
QY 559 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSVSPROCKL 618
Db 602 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSVSPROCKL 660
QY 619 ICRANGTGYFVVLAPKVVDTLCSPDSTSVQVQKCIKAGCDGNLGSKKRFDKCGVCGGD 678
Db 661 ICQAKGIGYFVVLQPKVVDTLCSPDSTSVQVQKCIKAGCDGNLGSKKRFDKCGVCGGD 720
QY 679 NKSCCKVTGLTKPMHGYNFVVAIPAGASSIDIRQYKGLIGDDNYLALKNQSGKYLIN 738
Db 721 GSTCKKISGVSATKPGYHDIITPTGATNIEVKQRNQRNNGSFLAIKAADGTIYL 780
QY 739 GHFVSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLETVEVLSVGKMTPPRVYSF 798
Db 781 GDYTLSTLEQDIYKGVLYRSGSSAALERIRSFSPKLEPLTQVLTGVNALRPKIKTY 840
QY 799 YLPKEPREDKSHHPKDPGPFVLSLNSVLSNOVEQDPRPARVAGSWGPCSCGSG 858
Db 841 FVKKK---KES-----FNAIPTFS-----AWVIEWGECSKSCSELG 873
QY 859 LQRAVDCRSAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCGGRGQR 915
Db 874 WQRLVECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCKTCGKGYK 930
QY 916 RSLKCVGHGGRLLARDQCNLHRRKPOE-LDFCVLRPC 950
Db 931 RSLKCLSHDGVLSHESCDPLKKPHFIDFCTMAEC 966

RESULT 8

US-60-453-050-14843
; Sequence 14843, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453, 050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14843
; LENGTH: 967
; TYPE: PRT

; ORGANISM: Homo sapiens
US-60-453-050-14843
Query Match 48.2% Score 2486.5; DB 7; Length 967;
Best Local Similarity 48.5% Pred No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;
QY 1 MLLGIITLAFAGTAGGSEPEREVVPIRLDPDINGRRYVWRGPDGGLQIFQITAF 60
Db 36 LLLAAALAVSDALGRPSEDEELVVP-ELE-----RAP-----GHGTTLRRLHAF 81
QY 61 QEDFYHLHTPDAQFLAFAPAFSTEHLG-----VPLQGLTGGSSDLRRFCYSGDVNAEPDSF 114
Db 82 DQQLDLRLPDSFLAPGFTLQNYGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAVSLGGLRGAFCAGYAEYVISPLPNAS---APAAQNSOGA---HLQRRGVPGGGS 167
Db 137 AALSCEGVRGAFYLLGEAFYIQLPAAERLATAAAGEKPPAPLOPHLLRR-----NRQ 191
QY 168 GDTSRGCV-----ASGNPAILRALDPYKPRRAGFESR 202
Db 192 GVGCGTGVVDDPRPTGKAETEDEGTEDEGAQWS-----PODPALQGVGP- 242
QY 203 SRRSGRAKRFVSIPIRYVETLVVADESVMKFGADLEHLLTLLATAARLYRHPSILNPI 262
Db 243 TGTGSIIRKRRFVSSHRYVETMLVADQSMABFHSGGLKHYLLTLFVSAARLYKHPSIRNSV 302
QY 263 NIIVVVKLLLRDRSDGPKVTGNAALTLRNFCAWQKLNKVSQKHPEYWDPTAILFTRODLC 322
Db 303 SLVVVKILVHDQKPEVTSNAALTLRNFCAWQKLNKVSQKHPEYWDPTAILFTRODLC 362
QY 323 GATTCDTLGMADVTCMDPKRSCSVIEDDGLSAFTTAHELGHVFNMPHNVKVEEVEFG 382
Db 363 GSOTCDTLGMADVTCMDPKRSCSVIEDDGLSAFTTAHELGHVFNMPHNVKVEEVEFG 422
QY 383 KLRANHMWSPTLQIDRANPWSACSAIITDFLDGSHGDCLLDOPSKPISLPEDLPASV 442
Db 423 VNQDSHMMASMLNLDHSQWSPSCSAIITDFLDGSHGDCLLDOPSKPISLPEDLPASV 482
QY 443 TLSQCELAFGVSGKPCP-YMQYCTKLWCTGKAGOMVCQTRHFPWADGTSCEGKCLK 501
Db 483 DANROCOFTFGEDSKHCPDAASTCTLWCTGSGVLVQCTKHFPWADGTSCEGKWCIN 542
QY 502 GACVERHNLKH---RVDGSKAKWDYPGCSRTCCGGVQLARRQCTNPTPANGKVCYEGV 558
Db 543 GKCVNKTD-RKHEDTFFHSGWMMGPMWDCSRTCCGGVQYTMRECDNPVFNKGKCYEGK 601
QY 559 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSVSPROCKL 618
Db 602 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSVSPROCKL 660
QY 619 ICRANGTGYFVVLAPKVVDTLCSPDSTSVQVQKCIKAGCDGNLGSKKRFDKCGVCGGD 678
Db 661 ICQAKGIGYFVVLQPKVVDTLCSPDSTSVQVQKCIKAGCDGNLGSKKRFDKCGVCGGD 720
QY 679 NKSCCKVTGLTKPMHGYNFVVAIPAGASSIDIRQYKGLIGDDNYLALKNQSGKYLIN 738
Db 721 GSTCKKISGVSATKPGYHDIITPTGATNIEVKQRNQRNNGSFLAIKAADGTIYL 780
QY 739 GHFVSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLETVEVLSVGKMTPPRVYSF 798
Db 781 GDYTLSTLEQDIYKGVLYRSGSSAALERIRSFSPKLEPLTQVLTGVNALRPKIKTY 840
QY 799 YLPKEPREDKSHHPKDPGPFVLSLNSVLSNOVEQDPRPARVAGSWGPCSCGSG 858
Db 841 FVKKK---KES-----FNAIPTFS-----AWVIEWGECSKSCSELG 873
QY 859 LQRAVDCRSAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCGGRGQR 915
Db 874 WQRLVECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCKTCGKGYK 930
QY 916 RSLKCVGHGGRLLARDQCNLHRRKPOE-LDFCVLRPC 950
Db 931 RSLKCLSHDGVLSHESCDPLKKPHFIDFCTMAEC 966


```
Db 340 KILVIHDEQKGPVTSNAALTLRNFNCWQKQHNPPSDRDAEHYDTAILFTIRQDLGSGQC 399
QY 328 DTLMADVGTMCDDPKKSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVEEVFGKLRAN 387
Db 400 DTLMADVGTVCDDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDNAKOCASLNGVNOQS 459
QY 388 HMMSPFTLIODRANPWSACSAATITDPLDSGHCDCILLDQPSKPISTLPEDLPASYSYLSQ 447
Db 460 HMASMLSLNDSQSPSCSAIYMITSLFNGHGECLMDKPNQPIQLPGDLPGTSDANRQ 519
QY 448 CELAFGVGSKPCP-YMQYCTKLWCTGAKQGMVQCOTRHPFPWADGTSCGEGKCLKGACVE 506
Db 520 CQFTFGEDSKHCPDAASTCTLWCTGSGVLVCQTKHPFWALGTSOGEGKWCINGKCVN 579
QY 507 RHLNLKH---RVDGSNAKWDYPGCSRTCGGGVQLARRQCTNTPANGGKAYCEGVRYKYR 563
Db 580 KTD-RKHFTDPTFHSGMGMGMPWDCSRTCGGVQYTMRECDNPVPKNGGKYCBGKRVYR 638
QY 564 SCNLEPCPSASCKSPREOCEAFNGYNHSTNRLTLAVAVWPKYSGVSPRDKCKLICRAN 623
Db 639 SCNLEDCPDN-NGKTFREEOCEAHNEFSKASFGSGGPAVEWIPKYAGVSPKDRCKLICQAK 697
QY 624 GTGYFVLAPKVVDGTLCSPDSTSVCGQKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCK 683
Db 698 GIGYFVLQPVVDGTPCSPDSTSVCGQCYKAGCDRIIDSKKFDKCGVCGGNGSTCK 757
QY 684 KYTGLETKPMHGNVFNVAIPAGASSIDIRQGYKGLIGDDNYALNKSQKYLINGHFVV 743
Db 758 KISGVSATKAPGVDHITTIPTGATNIEVKORNGSRNNGSFLAIRAADSTYLINGDYTL 817
QY 744 SAVERDLVKGSLRYSGTCTAVESLOASRPILLEPLTVEVLSVCKMTPPVRVYSFYLPR 803
Db 818 STLEQDMTKGVVLYRYSGSSAALERSFSPLKEPTLIQVLTGVNALRPKIKTYFVKKK 877
QY 804 PREDKSHHPDRGPRSVLNSVLSLNOVEQPDPRPPARVAGSWGPCASGSGSLQKRA 863
Db 878 ---KES-----FNAIPTFS-----ANVIEEWGECSSKCELGQORRL 910
QY 864 VDCRGSAQORTVPACDAAH---RPVETOACGE-PCPTWELSAWSPCKSGKSGRGQRSLKC 920
Db 911 VECDRINGQ---PASECAKEVKPASTRPCADHPCPOWOLGWSGSKTQCKGYKRSKLC 967
QY 921 VHGGRLLARDONLHRKPQE-LDFCVLRPC 950
Db 968 LSHDGGVLSHESCDPLKKPKHFDICTMARC 998

RESULT 11
US-10-369-779-29
; Sequence 29, Application us/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Sharma, Rahul
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Westphal, Ryan
; APPLICANT: Feder, John
; TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
; TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
; FILE OF INVENTION: ADAM-TS-SI
; FILE REFERENCE: D0204 NP
; CURRENT APPLICATION NUMBER: US/10/369,779
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/358,151
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-369-779-29
```

Query Match

48.1%; Score 2485.5; DB 6; Length 951;

RESULT 12

```
Best Local Similarity 48.0%; Pred. No. 2.7e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
QY 1 MLLLGITLTAFACTAGG--SEPEREVVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57
Db 20 LLLLASITMLLCARGAHGRPTDEELVLP-SLE-----RAPGHDSTTIRL--FL 66
QY 58 TAPQEDFYLHLTPDAQFLAPASTEHLGVLPLQGLTGS-----SDLRRCFYSGDVNA 109
Db 67 DAFGQQLHLKLODGSGLAPFTLTQV-----GRSPGSEAQHLDPDGTGLAHCFYSGTNG 121
QY 110 EPDSFAVSLCGRLGAFYVGAEEYVISLPNAS-----APAAQRNSOGA-----HLQRRG 161
Db 122 DPGSAAALSLCEGVGAFYLOGEEFFIQAPGVATRELPAVPEEBSARPQHILRRR- 180
QY 162 VPGSPGSDPTSRGCVASGNWNPAILRALDPYKPRRAGFEGESRRR----- 206
Db 181 ----RRSGGAKCGVMD-----DETLPDTSRSPESQNTNRNQWPVRDPTFODAGKP 226
QY 207 ----SGRAKRFYSIPRYVETLVVADESVMKFKGADLEHYLLTLATAARLYRHPSTILNFI 262
Db 227 SGPGSIRKKREFVSSPRIVETMLVADQSMADFHGSLKHLKLLTLFSAARFYKHPSTIRNSI 286
QY 263 NIWWWVLLLRDRDSGPKVTGNAALTLRNFCAWQKLNKVSDBKHPYWDFTAILFTQDLC 322
Db 287 SLVVVAILVYEBQKQPEVTSNAALTLRNFCAWQKHNQSPSDRDRPEHYDTAILFTQDLC 346
QY 323 GAPTCDTLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEVFG 382
Db 347 GSHTCDTLGMADVTGMDPKRSCSVIEDDGLQAAFTTAHELGHVFNMPHDNAKHCASLAG 406
QY 383 KLRANHMSPFTLIQDRANPWSACSAIITDPLDSGHCDCILLDQPSKPISTLPEDLPASYS 442
Db 407 VTGDSHLMASMLSSLDHSPQSPCSAYMTVSTFLDNGHGECLMDKPNQIKLPSDLPGTLY 466
QY 443 TLSQCELAFGVGSKPCP-YMQYCTKLWCTGKAKGOMVQCOTRHPFPWADGTSCGEGKCLK 501
Db 467 DANRQCFTEGEEKKICPDAASTCTTLWCTGSGGLVLCQTKHPFPWADGTSCGEGKWCVS 526
QY 502 GACVERHNLNKH---RVDGSNAKWDYPGCSRTCGGGVQLARRQCTNTPANGKYYCEGV 558
Db 527 GKCVNKTDM-KHFATPVHGSWGPWGPCSRTCGGVQYTMRECDNPVPKNGKYYCEGK 585
QY 559 RVKYRSCNLEPCPSASGSKSPREOCEAFNGYNHSTNRLTLAVAVWPKYSGVSPRDKCKL 618
Db 586 RVYRSCNLEDCPDN-NGKTFREEOCEAHNEFSKASFGNEPTVEWTPKTAGVSPKDRCKL 644
QY 619 ICRANGTYFYVLAPKVVDGTLCSPDSTSVCGQKCIKAGCDGNLGSKKRFDKCGVCGGD 678
Db 645 TCEAKGIGYFVVLQPKVVDGTPCSPDSTSVCGQCYKAGCDRIIDSKKFDKCGVCGGN 704
QY 679 NKSCKKVTGLTKPMHGNVFNVAIPAGASSIDIRQGYKGLIGDDNYALNKSQKYLIN 738
Db 705 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHNRQGRNNGSFLAIRAADGTYLIN 764
QY 739 GHVWFAVERDLVVKSLRYSGTCTAVESLOASRPILLEPLTVEVLSVCKMTPPVRVYSF 798
Db 765 GNPTLSLEQDLTYKTVLRYSGSSAALERSFSPLKEPTLIQVLMVGHALRPKIKFTY 824
QY 799 YLPKPREDKSSHHPKDRGPRSVLNSVLSLNOVEQPDPRPPARVAGSWGPCASGSG 858
Db 825 FMKKKTES-----FNAIPTFS-----EWVIEEWGECCKTCGSG 857
QY 859 LQKRAVDCRGSAQORTVPACDAAH---RPVETOACGE-PCPTWELSAWSPCKSGRGQR 915
Db 858 WQRRVVQCRDINGH---PASECAKEVKPASTRPCADLPCHPMQVGDWSPGSCSKTCGKYYK 914
QY 916 RSLKCVCHGGRLLARDONLHRKPQE-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 950
```



```
US-10-381-793-3
; Sequence 3, Application US/10381793
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN ADAM-TS-LIKE PROTEIN
; FILE REFERENCE: L10152 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/381,793
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: US 60/235,881
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/XXX,XXX
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-381-793-3

Query Match      48.1%; Score 2485.5; DB 6; Length 951;
Best Local Similarity 48.0%; Pred. No. 2.7e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

QY 1 MLLGILTLAFAGRTAGG--SEPERVWVPIRLDPDINGRRYWRGP-EDSGDQGLIFQI 57
Db 20 LLLASITMLLCARGAHRPTDEELVLP-SLE-----RAPHGDSITTRL--RL 66
QY 58 TAFQEDFYLHTPDQAFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA 109
Db 67 DAFGQQLHLKLPDGSGLFAPGFTLTQV-----GRSPGSEAQHLDPDGLAHCFYSGTVNG 121
QY 110 EPDSFAAVSLCGGLRGAFGYGAEVISPLPNAS-----APAAQRNSDGA-----HLLORRG 161
Db 122 DPGSAALSLCEGVGAYGAYLOGEFFIQAPGVATERLAPAVPEESSAFVQFHILRRR- 180
QY 162 VPGGSGDPTSRGCVASGWNPAIRLADPYKPRRAGFGEESRRR----- 206
Db 181 -----RSGSGAKCGVMD-----DETLPDSDRSPESQNTNRQWPKRPTQDAGKP 226
QY 207 ----SGRAKRFVSPRYVETLVVADESVMKFGADLEHYLLTLATARLYRHPISILNPI 262
Db 227 SGPGSIRKRFVSPRYVETLVVADESVMKFGADLEHYLLTLATARLYRHPISILNPI 286
QY 263 NIWVKVLLLRDRSDGPKVTGNAALTRNFCAMQKLNKVSQKHPYWDIAITLFTRODLC 322
Db 287 SLVVKVILVIEEQKGPVTSNAALTRNFCAMQKLNKVSQKHPYWDIAITLFTRODLC 346
QY 323 GATTCDTLGMADVGTMCDDPKRSCSVIEDDGLPSTAFTHAHLGHVFNPNPHDNPVKYCEVFG 382
Db 347 GSHTCDTLGMADVGTMCDDPKRSCSVIEDDGLPSTAFTHAHLGHVFNPNPHDNPVKYCEVFG 406
QY 383 KLRANHMSPPTLQIDRANPWSACSAALITDPLDSHGDCCLLDQPSKPISIPEDLPASV 442
Db 407 VTGDSHLMASMLSLDHSQWSPSCAYMVTSFLDNGHGECUMDRPKQPIKLPDLPGLTY 466
QY 443 TLSQOCELAFAFGSKPCP-YMOYCTKLWCTGKAKGQWVCQTRHPWADGTSCEGSKCLCK 501
Db 467 DANROCOFTFGEESKPCPDAASTCTTLWCTGSGLLVCQTKHFPWADGTSCEGSKWCVS 526
QY 502 GACVERHNLNKH---RVDGSAKWDPYGPCSRTCCGGVQLARQCTNPTFANGKCYCEGV 558
Db 527 GKCVKNTDM-KHFAATPVHSGMPGWDGDCSRTCCGGVQVYTRRCDNPNVKNQKCYCEG 585
QY 559 RVKRYSCNLEPCPSASGKSPRECEQAFNGYNHSTNRLLTAVAWVPKYSVSPRDKCKL 618
Db 586 RVRYSCNIEDCPDN-NKFTFRECEAHNEFSKASGNEPTVETWTPKYAGVSPDKCKL 644
QY 619 ICRANGTYFYVLAPKVVVDGTLCSPDSTSVQVQCKICAKGDCNGLSKKRFDRKCGVCGGD 678
Db 645 TCEARGIGYFVLQPKVVDGTPCSPDSTSVQVQCKVACGDCRIIDSKKAFDRKCGVCGGN 704
QY 679 NKSCKKVTGLTKPMHGYNFVAIPAGASSIDIRQRYKGLIGDNDYLAALKNSQKGYLLN 738
```

```
Db 705 GSTCKRMSGIVTSTRPGYHDIIVIPAGATNIEVKHRNQRNGSFLAIRADGTYILN 764
QY 739 GHFVVSAYERDLVVKVGLLRYSQGTAVESLQASRPILPLETLTVEVLVSGKMTPPRVYSE 798
Db 765 GNFTLSLTLEQDLTKYGVLRYSGSSAALERISFSLKEPLTTQVLMVGHALRPKIKFTY 824
QY 799 YLPKEPREDKSSHPKDPGRPSVLHNSVLSISNQVEOPDDRRPPARWVAGSNGPCSGSG 858
Db 825 FMKKKTES-----FNAIPTFS-----EWTEENGECSTKCGSG 857
QY 859 LOKRAVDCRGSAQORTVPCADAAH--RPVETOAGGE--PCPTWELSAWSPCKSKCGRCFOR 915
Db 858 WQRVVOCRDINGH---PASECAKEKVPASTRPCADLPCHPQVQVGDWSPCKSKCGKGYKK 914
QY 916 RSLKCVGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCPLKPKKHVDFCTLTQC 950
```

```
RESULT 13
US-10-369-779-19
; Sequence 19, Application US/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Sharma, Rahul
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Westphal, Ryan
; APPLICANT: Feder, John
; TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
; TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
; FILE REFERENCE: D0204 NP
; CURRENT APPLICATION NUMBER: US/10/369,779
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/358,151
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-369-779-19
```

```
Query Match      48.1%; Score 2485.5; DB 6; Length 968;
Best Local Similarity 48.0%; Pred. No. 2.8e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
```

```
QY 1 MLLGILTLAFAGRTAGG--SEPERVWVPIRLDPDINGRRYWRGP-EDSGDQGLIFQI 57
Db 37 LLLASITMLLCARGAHRPTDEELVLP-SLE-----RAPHGDSITTRL--RL 83
QY 58 TAFQEDFYLHTPDQAFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA 109
Db 84 DAFGQQLHLKLPDGSGLFAPGFTLTQV-----GRSPGSEAQHLDPDGLAHCFYSGTVNG 138
QY 110 EPDSFAAVSLCGGLRGAFGYGAEVISPLPNAS-----APAAQRNSDGA-----HLLORRG 161
Db 139 DPGSAALSLCEGVGAYGAYLOGEFFIQAPGVATERLAPAVPEESSAFVQFHILRRR- 197
QY 162 VPGGSGDPTSRGCVASGWNPAIRLADPYKPRRAGFGEESRRR----- 206
Db 198 -----RSGSGAKCGVMD-----DETLPDSDRSPESQNTNRQWPKRPTQDAGKP 243
QY 207 ----SGRAKRFVSPRYVETLVVADESVMKFGADLEHYLLTLATARLYRHPISILNPI 262
Db 244 SGPGSIRKRFVSPRYVETLVVADESVMKFGADLEHYLLTLATARLYRHPISILNPI 303
QY 263 NIWVKVLLLRDRSDGPKVTGNAALTRNFCAMQKLNKVSQKHPYWDIAITLFTRODLC 322
Db 304 SLVVKVILVIEEQKGPVTSNAALTRNFCAMQKLNKVSQKHPYWDIAITLFTRODLC 363
```

```
QY 323 GATTCTDLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKNVCEVEFG 382
Db 364 GSHTCTDLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKNVCEVEFG 423
QY 383 KLRANHMSPTLIQIDRANPWSACSAAITDLDGSHGDCLLDQSPKPSILPELPGASY 442
Db 424 VTGDSHLMASMLSSLDHSPQSPSAYWVTSFLDNGHGECLMDKPNPIKLPSPDLPGTLY 483
QY 443 TLSQOCELAFGVGSKPCP-YMQYCTKLWCTGKAGQWVQCTRHFPWADGTSCEGKCLCK 501
Db 484 DANRQCQFTTGEEBKPCDPAASTCTTLWCTGSGLLVQCTKHFPWADGTSCEGKWCVS 543
QY 502 GACVERHNLNKH---RVDGSNAKWDPYPCSRCTCGGVQGLARRQCTNPTFANGKYCEGV 558
Db 544 GKCVNKTDM-KHFATPVHGSWGPWGDCSRTCGGVQYTMKCDNPNVKNNGKYCEGK 602
QY 559 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAVVPKYSGVSPRDKCKL 618
Db 603 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAVVPKYSGVSPRDKCKL 661
QY 619 ICRANGTGYFVVLAPKVPVDTGLCSPDSTSVCGQKICAKGDCNGLGSKKRFKDCGVCGD 678
Db 662 TCEAKGIGYFFVLOPKVPVDTGLCSPDSTSVCGQKICAKGDCNGLGSKKRFKDCGVCGD 721
QY 679 NKSKCKVTGLFTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNYLALXNSGKYLIN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHNRQGRNNGSFLAIRAADGTIYL 781
QY 739 GHEVSAVERDLVVKSLRLYSTGTAVESLQASRPLEPLVEVLSVKGKMTPPRVYSF 798
Db 782 GNTLTSLTLEODLYKTVLRYSGSSAALIRSFSPKREPLTIQVLWGHALRPKIKETY 841
QY 799 YLPEKREDKSHPKDPRGSPVLHNSVLNSQVEQDPRPARVAGSWGKPCSCSGSG 858
Db 842 FMKKTES-----FNAITFS-----ENVIEWGECSKTCGSG 874
QY 859 LQKRAVDCRGSAGQRTVPACDAH--RPVETOACGP-PCPTWELSAWSPCSKSGRGQR 915
Db 875 WQRRVQCRDINGH---PASECAKEYKPASTRCPADLPCHPWQGVGDSWSP-SKTCGKGYK 931
QY 916 RSLKCVGHGRLARQCNLHRRPOB-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDLPKPKHYIDFCILTQC 967

RESULT 14
US-10-391-364-82
; Sequence 82, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16558, 14223, 16002, 50566, 50552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MPI03-0190NMN
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
```

```
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper of PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-391-364-82
```

```
Query Match 48.1%; Score 2485.5; DB 6; Length 968;
Best Local Similarity 48.0%; Pred. No. 2.8e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
```

```
QY 1 MLLLGILTLAFAGTAGG--SEPEREVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFCI 57
Db 37 LLLASITMLLCARGAUGRPTDEELVLP-SLE-----RAPHDSITTRL--RL 83
QY 58 TAFEDEFYHLTPDAQFAPAFSTEHLGVPLQGLTGS-----SDLRRCFYSGDVNA 109
Db 84 DAFGQQLHLKLPDPSGLFAPGFTLQTV-----GRSPSEAQLHDPDGLAHCFYSGTVNG 138
...QY 110 EPDFAAVSLCGGLRGAFFGYRGAIVYISPLPNAS----APAAQNSOGA---HLLQRRG 161
Db 139 DPGSAALSLCEGYRGAFFYLOGEEFFTOPAPGVATEKLAPAVEEESARPQHILRRR- 197
QY 162 VPGPGSDPTSRGCVASGWNPAILRALDPYKPRRAGGESRRR----- 206
Db 198 -----RRSGGAKCGVMD-----DETLPSTDSRPSQNTNRQWVRDPTPDAGKP 243
QY 207 ----SGRAKRFVSTPRVETVLVADESMVAFPHGADLHLYLLTLATAARLYRHPISILNFI 262
Db 244 SGPGSIRKKRFVSPRYVETMLVADQSMADPHGSLGKHYLLTLFSAARFYKHPISIRSI 303
QY 263 NIVVVKVLLLRDRDSGPKVTGNAALTIRNFCAMOKKLVSKDKHPEYWDTAILFTRODLC 322
Db 304 SLVVVKILVIYEEQKGPEVTSNAALTIRNFCNWKQINSFSDRDEHYDTAILFTRODLC 363
QY 323 GATTCTDLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKNVCEVEFG 382
Db 364 GSHTCTDLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKNVCEVEFG 423
QY 383 KLRANHMSPTLIQIDRANPWSACSAAITDLDGSHGDCLLDQSPKPSILPELPGASY 442
Db 424 VTGDSHLMASMLSSLDHSPQSPSAYWVTSFLDNGHGECLMDKPNPIKLPSPDLPGTLY 483
QY 443 TLSQOCELAFGVGSKPCP-YMQYCTKLWCTGKAGQWVQCTRHFPWADGTSCEGKCLCK 501
Db 484 DANRQCQFTTGEEBKPCDPAASTCTTLWCTGSGLLVQCTKHFPWADGTSCEGKWCVS 543
QY 502 GACVERHNLNKH---RVDGSNAKWDPYPCSRCTCGGVQGLARRQCTNPTFANGKYCEGV 558
Db 544 GKCVNKTDM-KHFATPVHGSWGPWGDCSRTCGGVQYTMKCDNPNVKNNGKYCEGK 602
QY 559 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAVVPKYSGVSPRDKCKL 618
Db 603 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAVVPKYSGVSPRDKCKL 661
QY 619 ICRANGTGYFVVLAPKVPVDTGLCSPDSTSVCGQKICAKGDCNGLGSKKRFKDCGVCGD 678
Db 662 TCEAKGIGYFFVLOPKVPVDTGLCSPDSTSVCGQKICAKGDCNGLGSKKRFKDCGVCGD 721
QY 679 NKSKCKVTGLFTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNYLALXNSGKYLIN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHNRQGRNNGSFLAIRAADGTIYL 781
```

QY 739 GHFVSAVERDLVWKSLLRYSGTAVESLQASRPILPTVEVLSVG4MTPPRVRSF 798
Db 782 GNFTLSTLEODLTYAGTVLRYSGSAALEIRSFSPLEPTLQVLMVGHALRPKIRFTY 841
QY 799 YLPKPREDKSHPKDPRGSPVLHNSVLSLNOVEQDDRPAPRWAGSNGPCSASGSG 858
Db 842 FMKKKTES-----FNAIPTFS-----BWWLEWGECSKTGSG 874
QY 859 LQKRAVDCRSGAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSGRQFOR 915
Db 875 WORRIVQCRDINGH---PASECAKEVKPASTRCPADLPCHWQVGDWSPCSKTCGKYKK 931
QY 916 RSLKCVGHGRLRLADQCNLHRKPQE-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 967

RESULT 15
US-60-440-068-134
; Sequence 134, Application US/60440068
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CARMAN, JULIE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE
; TITLE OF INVENTION: NF-KB PATHWAY
; FILE REFERENCE: 3053-4191
; CURRENT APPLICATION NUMBER: US/60/440,068
; NUMBER OF SEQ ID NOS: 2003-01-14
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 134
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-440-068-134

Query Match 48.1%; Score 2482.5; DB 7; Length 967;
Best Local Similarity 48.9%; Pred. No. 4.6e-174;
Matches 485; Conservative 155; Mismatches 250; Indels 101; Gaps 24;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSGQGLIFQITAF 60
Db 36 LLLLAALLAVSDALGRPSEDELVVP-ELE-----RAP---GHGTRRLRLHAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGSSDLRRCFYSGDVNAEPDSF 114
Db 82 DQQLDLRLPDSFLAPGFTLQVGRKSGSETPLP-----ETDLAHCFTSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFGYGAAYVISPLPNAS---APAAQNSQGA---HLLQ---RRGYPGG 165
Db 137 AALSICEVGRGAFYLLGEAVFIQPLPAASERLATAAPGEKPPAPLQPHLLRRNRQGDVG 196
QY 166 PSG-----DPTSRC-----GVASG--WNPAIRLALDPYKPRRAGFSGSRSSRS 207
Db 197 TCGVVDEPRPTGKAETDEDEGEDEGPQWS-----PQDPALQGVQP-TGTGS 247
QY 208 GRAKRFVSIPIRYVTLLVVADESVMKFGADLEHYLLTLATAARLYHHPSLINPINIVV 267
Db 248 IRKRFVSSHRYVTMLVAQSMAEFGSLGKHYLLTLFVAARLYKHPSIRNSVSLVV 307
QY 268 KVLRLDRDSGPKVTGNAALTLRNFCWQKLNKVDKHPYWDTAILFTRODLGATTC 327
Db 308 KILVIHQKGPVTSNAALTLRNFCWQKLNKVDKHPYWDTAILFTRODLGATTC 367
QY 328 DTLGNADYGVWCDPKRSCSVIEDGLPSAFTTAHELGHVFNMPHDNVKVEEYFGKLRAN 387
Db 368 DTLGNADYGVWCDPKRSCSVIEDGLPSAFTTAHELGHVFNMPHDNVKVEEYFGKLRAN 427
QY 388 HMSPTLIQIDRANPWSACSAIITDLDGSHGDCILDDQPSKPLSLPEDLPAGASYTLQQ 447
Db 428 HMASMLSNLDHSPWSPCSAYMITSLDNGHGECLMDKPHNPIQLPGLDLPGSTYDANRQ 487

QY 448 CELAFGVGSKPCP-YMOYCTKLACTGKAKQVMQVQTRHFPWADGTSCGSGKLCCLKACVE 506
Db 488 CQTFGEDSKHCPDAASTCTLWCTGTSGGLVLCQTKHFPWADGTSCGSGKWCINGKCVN 547
QY 507 RHNLNKH--RVDSWAKWDPYGPCSRTCGGQVQLARRCCTNPTPANGKCYCGVVKYK 563
Db 548 KTD-RKHFDTPFHGNWGMWGPWGDSCRTCGGQVQYTWRECDNVPKNGKCYCGKVKYK 606
QY 564 SCNLEPCPSASGKSPREEOCEAFNGYNHSTNRLTLAVAWVPKYSGVSPDRCKLILCRAN 623
Db 607 SCNLEDCPDN-NGKTFREEOCEAHNEFSAFSGSPAVENIPKYAGVSPKDRCKLILCOAK 665
QY 624 GTGYEVVYLPKVVVDGTLCLSPDSTVQVQKCIKAGCDGNLGSKKRFDKCGVCGGJNKSCK 683
Db 666 GIGYFFVLQPKVVVDGTPCSPDSTVQVQKCVKAGCDRIIDSKKKFDKCGVCGGNGSTCK 725
QY 684 KVTGLTKPMHGYNFVVAIPAGASSIDIRQYKGLIGDNDYLLALKNSSQKYLNLNGHFVV 743
Db 726 KISGVSITSAKPGYHDIITPTGATNIEVKQNRQGRNRNGSFLAIKADGTIYLNQDYL 785
QY 744 SAYERDLVYKSLRLRYSGTAVESLQASRPILPTVEVLSVGKMTTPRVRSFYLPKE 803
Db 786 STLEQDIMYKVVLYRYSGSAALEIRSFSPLEPTLQVLTGVGNALRPKIRTYEVKKK 845
QY 804 PREDKSHPKDPRGSPVLHNSVLSLNOVEQDDRPAPRWAGSNGPCSASGSGJLQKRA 863
Db 846 -----KES-----FNAIPTFS-----AWVTEWGECSKSELGMQRRL 878
QY 864 VDCRGSAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSGRQFORSLKRC 920
Db 879 VECRDINGQ---PASECAKEVKPASTRCPADHPCPQWQLGSEWSSCSKTCGKGYKKSLAC 935
QY 921 VGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKKPKHYIDFCTMAEC 966

Search completed: May 9, 2003, 15:29:12
Job time : 102 secs

